Informing Youth Health and Well-Being Programs: A National 4-H Healthy Living Evaluation Study

Donna J. Peterson
Mississippi State University
donna.peterson@msstate.edu

Laura H. Downey
Mississippi State University
laura.downey@msstate.edu

JoAnne Leatherman
National 4-H Council
jleatherman@fourhcouncil.edu

Abstract

4-H Healthy Living programs address healthy eating; physical activity; social-emotional health and well-being; alcohol, tobacco, and other drug use prevention; and injury prevention. Using the Systematic Screening and Assessment Method, this paper identified 32 4-H Healthy Living programs across the nation ready for comprehensive outcome evaluation and/or national replication based on 6 criteria. Weaknesses in an additional 78 programs that did not meet the criteria were also identified. Programs that failed to meet the criteria did so primarily because they lacked a clearly delineated theory of change or appropriate evaluation. Implications for practice include ways to strengthen program planning and use of a comprehensive evaluation framework. Specific attention is given to professional development for 4-H professionals.

Key words: healthy living, 4-H, effective programs, environmental scan, evaluability assessment

Introduction

4-H is the largest youth development organization in the United States. In 2016, nearly 6 million youth ages 5 to18 participated in programming, and approximately 600,000 adults and youth

served as volunteers (National 4-H Council, 2016). Although health has been an integral part of 4 H since the early 20th century, the national 4-H Healthy Living initiative began in 2008.

4-H Healthy Living programs use a holistic approach that includes eating a healthy diet, engaging in safe physical activity, avoiding risky behavior, recognizing and directing emotions, and developing and maintaining positive social interactions and relationships (National 4 H Healthy Living Task Force, 2009). This holistic approach helps youth and their families increase their awareness, knowledge, skills, and competencies in physical, social, and emotional well-being. Thus, 4-H Healthy Living programs address five domains: healthy eating; physical activity; social-emotional health and well-being; alcohol, tobacco, and other drug (ATOD) use prevention; and injury prevention. A guiding principle of 4-H Healthy Living programs is that programs and curricula are based on best practices in healthy living research (National 4-H Healthy Living Task Force, 2009).

Despite the noted success and longevity of 4-H, Arnold (2015) argued that 4-H does not consistently adhere to a practical program model "that articulates the 4-H program theory of change or chain of action to guide program development and implementation" (p. 55). Program theory is critical because it can identify steps needed to achieve desired outcomes and result in a blueprint for implementation and evaluation (Rennekamp & Jacobs, n.d.). Kettner, Moroney, and Martin (2017) suggest that a logical approach to programming is more likely to produce desired outcomes, as compared to simply hoping that a program achieves some outcome.

In an effort to build logical connections between 4-H programs and outcomes, National 4-H Council (2010) developed a logic model for its healthy living programs. However, leaders in 4-H were unsure of the degree to which programs adhered to that logic model, as well as whether the programs were achieving intended outcomes. This uncertainty prompted National 4-H Council to fund a project in 2013 to document quality 4-H Healthy Living programs (National 4-H Council, 2012).

Evaluation specialists with Mississippi State University Extension Service (MSU Extension) used the Systematic Screening and Assessment method (SSA) to identify promising 4-H Healthy Living programs that adhered to the National 4-H Healthy Living mission and logic model; met specific, minimal quality standards and criteria; and were deemed ready for comprehensive outcome evaluation and/or national replication. For this project, a program was defined as an organized, purposeful set of activities (National 4-H Council, 2010). Information was also gathered on the quality of those programs that did not meet the minimal standards. The

growing number of healthy living programs that target youth, along with minimal evidence regarding the effectiveness of these programs, provided a timely opportunity to use SSA and contribute to the practice-based evidence of healthy living programs for youth (Dawkins et al., 2010). This article describes the process and outcomes of this project and presents implications for practice.

Methods

Application of SSA has been described in detail by Downey, Peterson, Le Menestrel, Leatherman, and Lang (2015). SSA combines an environmental scan and evaluability assessment in one comprehensive process (Leviton & Gutman, 2010). In brief, SSA steps included the following:

1. Scan for programs related to the 4-H Healthy Living domains.
2. Review the identified programs using predetermined criteria to assess program quality.
3. Conduct an evaluability assessment of those programs that met the criteria.
4. Classify the programs based on documentation of outcomes through various evaluation designs.
5. Report the information to inform future program efforts.

When scanning for programs (Step 1), multiple techniques were used (questionnaire, interviews, content analysis). First, we developed a 28-item online questionnaire as one technique in scanning for programs (Peterson, Downey, Leatherman, Le Menestrel, & Lang, 2015). Representatives from National 4-H Council and 4-H National Headquarters reviewed the questionnaire. After the questionnaire was finalized, we invited State 4-H Healthy Living Liaisons and State 4-H program leaders to complete the questionnaire online or forward it to another individual who could provide the requested information. We solicited information on the following topics related to healthy living programming: program name, 4-H Healthy Living domain, program objectives, target audience, curricula used and program resources, program activities and delivery methods, desired outcomes, evaluation methods and results, and program website and primary contact. Participants could report this information for up to 15 programs. We followed Dillman's tailored design method (2007) for questionnaire implementation. At least one program was reported from at least one land-grant university in all Cooperative Extension regions (nine from the North Central region, 10 from the Northeast region, eight from the Southern region, 12 from the Western region, and seven from the 1890 region, which includes portions of the southern and eastern United States). We reviewed

responses as they were submitted and when additional information about a program was needed (e.g., information in the questionnaire was missing or unclear), we conducted interviews with the primary contact.

Simultaneously, we conducted a content analysis of 4-H Healthy Living documents to identify programs not reported through the questionnaire process. Documents reviewed included grantee reports of 4-H Healthy Living projects funded by Walmart, United Healthcare, and Coca Cola; the 4-H Programs of Distinction database; and the 4-H Healthy Living Literature Review (Hill, McGuire, Parker, & Sage, 2009). The reports were reviewed for information similar to that sought through the online questionnaire (e.g., program name, objectives, audience, evaluation).

We used the following criteria during this scanning process (Step 2) to assess each program submitted through a national survey and/or identified in a published 4-H Healthy Living programmatic report:

Criterion 1: 4-H Healthy Living categories/domain(s) were identified or identifiable.

Criterion 2: Target population identified and is in 9 to 19 year age range.

Criterion 3: Specified objectives that are clear, realistic, and measurable.

Criterion 4: Objectives, activities, and outcomes are logically connected.

Criterion 5: Clearly specified, measureable, and realistic outcomes that are tied to the 4 H Healthy Living logic model outcomes.

Criterion 6: At least a pretest/posttest is used to assess outcomes, and evaluation results are available and/or reported.

The scanning resulted in an enumeration of criteria unmet and identification of common problems among those programs that failed to meet all six criteria (see Findings section).

Programs that met all criteria were included in the evaluability assessment (Step 3) and we used a checklist for program evaluability (developed by the United Nations Development Fund for Women, 2009) to assess the following evaluability parameters: program design, availability of information, and conduciveness of context for evaluation. Once again, we obtained more information about a program (if needed) through interviews or email communications with a program representative.

In Step 4, we classified programs as having preliminary, moderate, or strong evidence of replicability as defined by the Corporation for National and Community Service (CNCS, 2012):

- Preliminary evidence: Programs demonstrated desired outcomes through a pretest/posttest study design;
- Moderate evidence: Programs demonstrated desired outcomes through at least one experimental/quasi-experimental study design, or correlational research that controlled for selection bias; or
- Strong evidence: Programs demonstrated desired outcomes through the use of numerous quasi-experimental studies, or one large, well-designed and well-implemented, randomized, controlled trial implemented over multiple sites.

Thus, 4-H Healthy Living programs were categorized by level of evidence to identify those ready for comprehensive outcome evaluation and/or replication.

Findings

Through SSA, we identified 110 unduplicated 4-H Healthy Living programs. Figure 1 displays the flow scheme for the SSA. The domains with the largest number of identified programs were physical activity and healthy eating (co-classified) ($n = 43$), followed by social-emotional health and well-being ($n = 15$), and injury prevention ($n = 14$).

Thirty-two of the identified programs showed readiness for further outcome evaluation or had evidence of potential for replicability; the majority were in the physical activity and healthy eating domains ($n = 14$), followed by social-emotional health and well-being ($n = 7$). The appendix lists these programs and their references, domain, level of evidence, and URL for additional information. Each of these programs met several indicators within the evaluability parameters (program design, availability of information, and conduciveness of context). For example, responses about the program identified a clearly defined problem, an appropriate target population, a clear and accurate logic model or description of the program's theory of change, clear and measurable objectives, results from at least a pre/post evaluation, and resources to undertake a more rigorous evaluation. Figure 1 also identifies the domains and evidence of replicability classifications for these programs.

Most programs with evidence of replicability had preliminary evidence ($n = 25$). For example, in the physical activity and healthy eating domain, one program with preliminary evidence had a clearly defined problem (unhealthy eating habits and lack of physical activity by youth), specific target population (elementary and middle school youth), complete logic model, published

Figure 1. Overall Flow Scheme for the Systematic Screening and Assessment
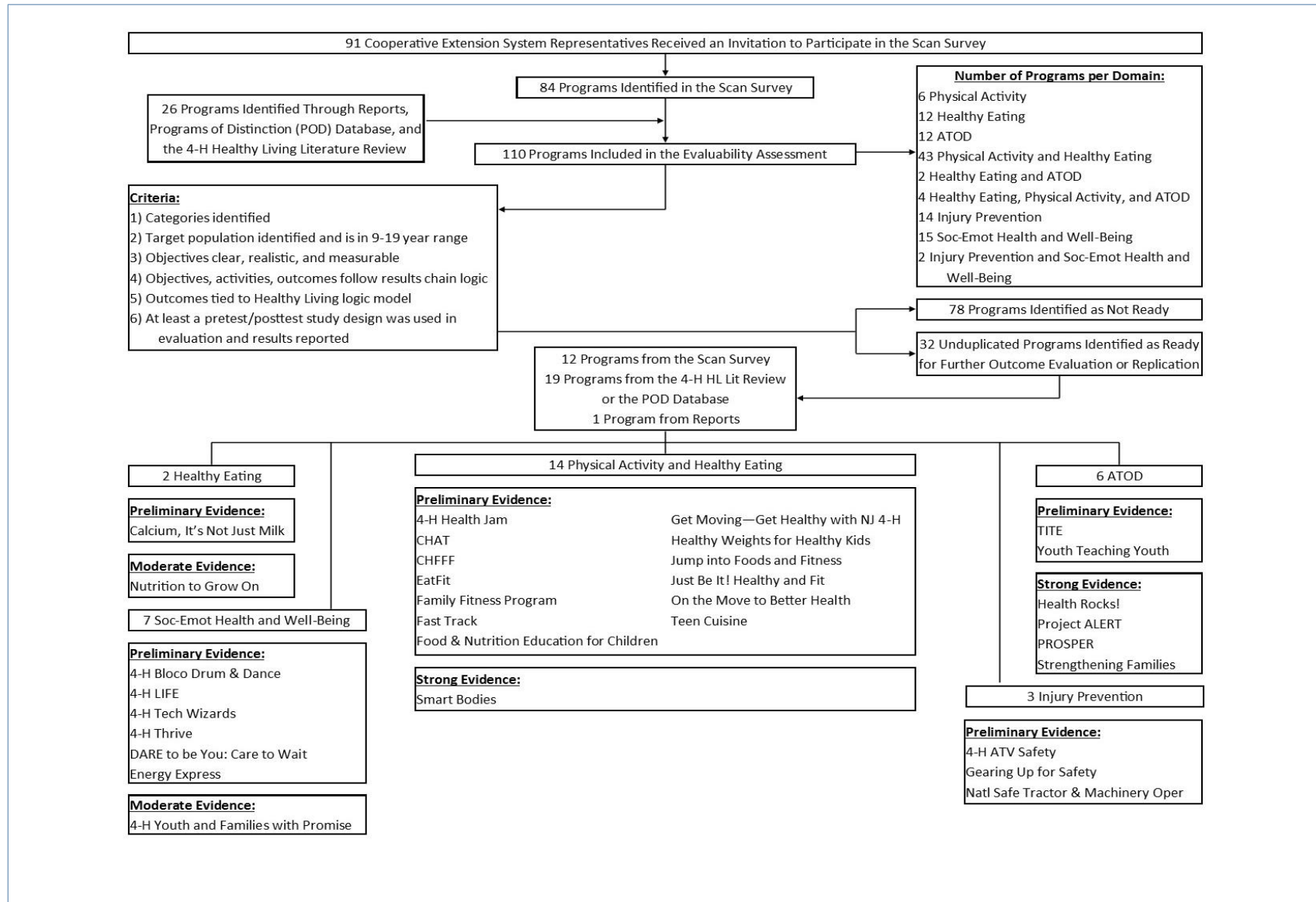
curriculum with appropriate activities, and resources for educators implementing it, but only pre/post-program outcome evaluation results.

We classified 78 of the 110 programs as "not ready for replication" because they failed to meet the six criteria (described previously). In numerous cases, multiple criteria were not met by a program. Table 1 documents the unmet screening criteria within each of the five domains.

All programs met criteria 1 and 2 (identifiable domain and target population within the 9 to 19 age range). Across each domain, programs most often lacked program logic or theory (criterion 4). For example, some programs' objectives were health related, but the only outcomes assessed were life skills. Although life skills can be an appropriate outcome, assessing *only* life skills will not provide evidence as to whether the specific health-related content changed health knowledge or behaviors. In other cases, programs described a desired outcome of increasing the amount of time spent doing physical activity, but opportunities for participants to practice physical activity were not included in the program.

Programs also often failed to use at least a pretest/posttest design (or retrospective pretest design) to evaluate program outcomes (criterion 6). Most programs that failed to meet criterion 6 used a posttest only or focused efforts on evaluating implementation of program activities. To document change in an outcome, at least two measurement points are needed. For example, if a program only administered a posttest, it was difficult to know if youth increased their health-related content knowledge because there was not a baseline assessment of where their health-related knowledge started.

Less frequently, program objectives and/or program outcomes were not clear, realistic, and measurable (criteria 3 and 5). Some program objectives were written as process objectives (what program staff would do) instead of outcome objectives (what participants would know or do as a result of the program). In one case, a program had the stated objective of "teach youth about healthy eating." In other cases, program outcomes were not realistic; for example, a program of only 3 hours duration indicated a desired outcome of a change in body mass index, which cannot occur in that short time frame.

Table 1. Programs with Unmet Screening Criteria within Each 4-H Healthy Living Program Domain

	Healthy eating	Physical activity	ATOD prevention	Social-emotional health and well-being	Injury prevention
Criterion 1: 4-H Healthy Living categories/domain(s) were identified or identifiable	0	0	0	0	0
Criterion 2: Target population identified and is in 9-19 year range	0	0	0	0	0
Criterion 3: Specified objectives that are clear, realistic, and measurable	9	7	0	3	3
Criterion 4: Objectives, activities, and outcomes are logically connected	22	19	1	4	4
Criterion 5: Clearly specified, measureable, and realistic outcomes that are tied to the 4-H Healthy Living logic model outcomes	8	3	0	4	4
Criterion 6: At least a pretest/posttest is used to assess outcomes and evaluation results are available and/or reported	12	12	4	11	6

Note. Some programs failed to meet multiple criteria, so a single program may contribute to counts in more than one cell in Table 1. For example, a program classified in the combined healthy eating and physical activity domain may have failed to specify clear, realistic, and measurable objectives (criterion 3) with respect to physical activity, and may also have failed to logically connect objectives (criterion 4) with respect to both healthy eating and physical activity, thus that one program contributes to the counts of 7, 22 and 19 in those cells in Table 1.

Discussion

The nationwide reach of 4-H and a focus on healthy living provided an excellent place to start with a systematic process of identifying programs that promote the health of youth. The extensiveness of programs also provided an opportunity to identify overall strengths and weaknesses in programming across 4-H for program improvement. Although 32 programs met the criteria for quality programming and evaluation, 78 programs did not. Two primary weaknesses were observed among those programs that were not identified as being ready for replication or more intensive outcome evaluation: lack of fully-delineated program theory and program evaluation approaches that could not adequately assess change resulting from the program.

A lack of logical connections among objectives, planned activities, and/or outcomes was a major weakness, indicating that a theory of change was likely not present. As mentioned, a theory of change can increase the chance that programmatic strategies chosen will accomplish desired outcomes. However, practitioners fail to develop this theory of change for many reasons. For example, in response to pressing needs or interests (e.g., an emerging community crisis or condition such as the opioid epidemic), 4-H campus or county professionals may rush to implementation, rather than thinking through whether program goals and objectives are well-defined and feasible, the change process is plausible, program procedures and activities are clear and adequate, and evaluation methods can document outcomes. The need to be responsive to needs while still addressing program theory has been discussed by Knowlton and Phillips (2013):

> Ideally, theories of change are grounded in literature, experience,
> or other evidence that promotes plausibility. If the theory of
> change is supported by a body of evidence, there is a stronger
> chance that the strategies chosen will secure the desired results.
> Frequently, however, this "standard" is overlooked. In the urgent
> fever to get to implementation, the design and plan quality can be
> shortchanged and rely, instead, on faulty assumptions, old
> practice, or little or no evidence (p. 17).

Creating a theory of change gives attention to the connections among programmatic components (Arnold, 2015). Developing a theory of change can occur through delineating a series of if-then relationships. In working out the if-then sequences, connections can be made,

assumptions can be clarified, and an understanding of how investments are likely to lead to results is enhanced. Figure 2 gives an example of how one might make these connections.

Figure 2. Example of Developing a Theory of Change through a Series of If-Then Relationships
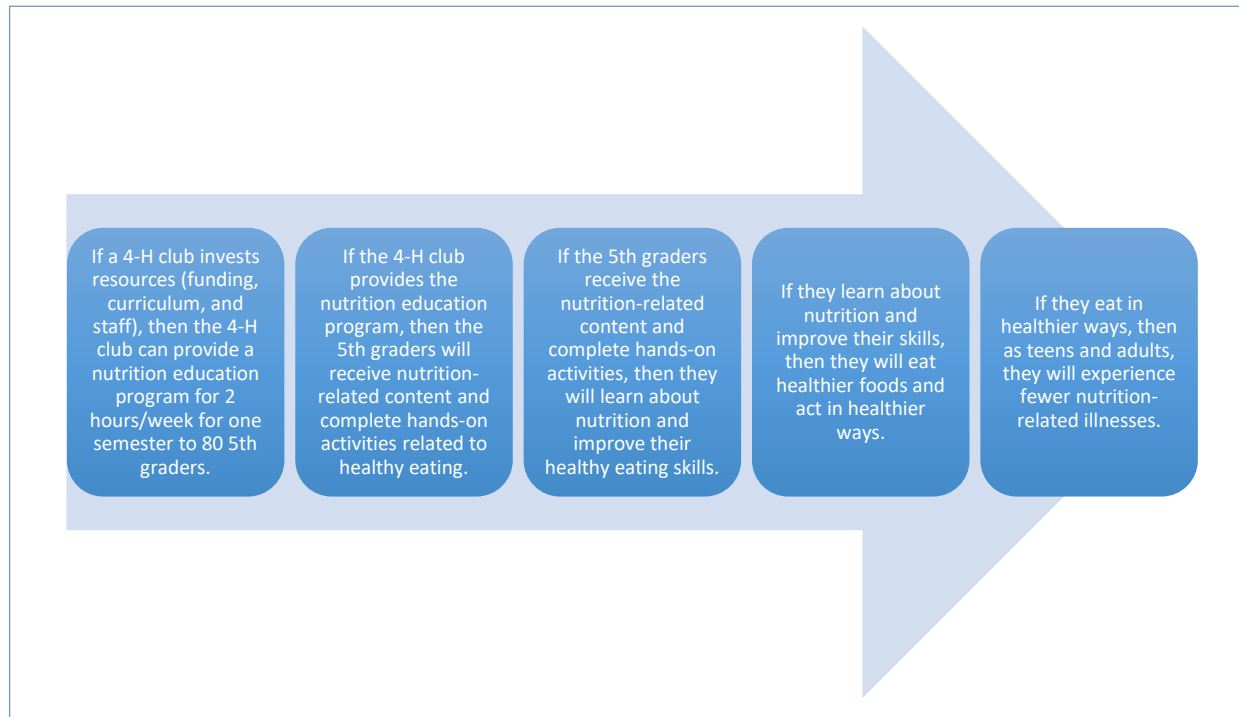
The second leading weakness was related to program evaluation. As mentioned earlier, programs that were not identified as being ready for replication or more intensive outcome evaluation often used a posttest only or focused exclusively on process or implementation evaluation. At least a pretest/posttest design that collects data from participants before and after a program is needed to estimate the program's effect by comparing data from these two measurement points (Rossi, Lipsey, & Freeman, 2004). Additionally, while process or implementation evaluation are important for helping identify reasons that a program failed to achieve its objectives, it does not provide evidence that the desired outcomes were met. Ideally, thorough evaluation would include both process or implementation evaluation and outcome evaluation using at least a pretest/posttest design.

The two weaknesses that were observed (i.e., lack of fully-delineated program theory and weak outcome evaluation design) could indicate broader challenges in Extension, and specifically 4-H.

While we have highlighted these weaknesses in this paper, approximately one-third of the programs reviewed for evaluability were classified as having preliminary, moderate, or strong evidence of readiness for replication at a national level or more rigorous outcome evaluation. In light of these challenges and strengths, implications for practice and recommendations follow.

Implications for Practice

This paper identified 32 4-H Healthy Living programs that are ready for further outcome evaluation or had evidence of potential for replicability on a national level. Implementation of programs with evidence of achieving desired outcomes can contribute to efficient use of programmatic resources. Thus, the next logical approach is to promote their use across 4-H systems or even other youth development organizations. 4-H programs are often easily accessible (e.g., low cost, collaborative nature of 4-H professionals). Additionally, those programs identified in this paper already have clear lessons and activities, identified evaluation tools and approaches, and supportive materials for educators. This would make adoption across 4-H, and even by other youth development organizations, attractive.

As an organization, 4-H is dedicated to ensuring that its programs are high quality, with specific objectives, activities, and outcomes that are clear, realistic, measurable, and logically connected, as well as having appropriate evaluation approaches to document outcomes achieved. As indicated by the National 4-H Healthy Living Task Force (2009), 4-H is dedicated to:

> Increasing the knowledge and commitment of Extension staff to design effective process evaluation strategies that enable newly developed learning experiences and curriculum to be improved, establish an ongoing monitoring process to ensure quality implementation, and create processes to eliminate and redirect resources away from ineffective programs.
> Designing effective evaluation strategies that enable 4-H professionals to develop healthy living curriculum to advance to the highest level of evidence possible.

This dedication includes attention to program theory and program evaluation—the two weaknesses reported in the current paper. This desire serves as a great strength and can be the driving force to promote professional development for 4-H educators and specialists. As the following examples will suggest, such professional development can help ensure that resources

dedicated to program development and implementation are not wasted due to a lack of program theory or weak program evaluation.

While Extension professionals may have strengths and expertise in certain content areas, working with youth, and/or communication, they may not have strengths or expertise in program development or program evaluation. Individuals who lack formal training in program development may be unable to clearly delineate a program's theory of change. However, even without expertise in those areas, beginning with a series of if-then statements to describe the components required to move from a program need to a desired outcome can be an initial step in developing a program's theory of change. Professional development related to theory of change models can be provided to help them learn how to integrate practice, experience, theory, and research to create a program that will be likely to achieve its desired outcomes. Without such integration, a program's content and/or activities may be highly innovative, but intended outcomes may be less likely to result (Knowlton & Phillips, 2013).

Without expertise in program evaluation, 4-H professionals would benefit from the use of a comprehensive evaluation framework that delineates components needed to thoroughly assess both a program's process and outcomes. The RE-AIM framework is one option that has been promoted for use in 4-H Healthy Living programs (National 4-H Healthy Living Task Force, 2009). RE-AIM (re-aim.org) identifies five evaluation dimensions (Reach, Effectiveness, Adoption, Implementation, and Maintenance) and has been successfully used to inform the selection of evidence-based health promotion programs (Glasgow, Vogt, & Boles, 1999). RE-AIM is a straightforward and consistent approach that could be beneficial in several ways to (a) assess a 4-H program's health outcomes, (b) compare program processes and health outcomes across multiple sites or over time, and (c) inform decisions about resource distribution for effective programs.

Given these benefits, Downey, Peterson, Donaldson, and Hardman (2017) applied RE-AIM to a 4-H Healthy Living program that was implemented in Mississippi and Tennessee. RE-AIM provided a standardized evaluation protocol that documented program participants or reach, implementation, adaptation, and outcomes across multiple sites in two states and diverse settings in which the program was implemented (e.g., school classrooms, community centers). Use of the RE-AIM framework ensured that multiple components needed for a comprehensive program evaluation were included, and thus helped avoid the weaknesses related to evaluation identified through the evaluability assessment reported here. For example, both process evaluation (Reach, Adoption, and Implementation) and outcome evaluation (Effectiveness and

Maintenance) were conducted, with measurement of the effectiveness and maintenance components requiring data collection at multiple time points to show true change.

Identification of these two main weakness and suggestions for addressing them are of great importance to administrators, educators, specialists, and volunteers in 4-H. Individuals who focus on professional development and capacity building can use this information to enhance the quality of the programs offered. Professional development could take the form of recorded online presentations, interactive webinars, face-to-face trainings, one-on-one technical assistance, or even printed materials, based on the resources available in the state 4-H system. Administrator support and encouragement for such training and technical assistance on program theory of change and/or RE-AIM (or another evaluation framework) would show that such knowledge and its use are valued. Additionally, administrators could use evaluation results from a consistent framework to compare programs in order to inform decisions about which programs to continue, expand, modify, or eliminate.

The MSU Extension evaluation specialists who led this project have begun such professional development efforts at a national level with different target audiences. For example, a competency building workshop for 4-H agents was conducted at the 2014 Annual Conference of the National Association of Extension 4-H Agents in Minneapolis, Minnesota. Similar workshops with Extension professionals, including those professionals who work primarily in program and staff development, were delivered at the 2013 Annual Conference of the American Evaluation Association in Washington, DC; at the 2016 Annual Conference of the National Association of Extension Program and Staff Development Professionals in Ridgedale, Missouri; and during the 2015 Evaluation Virtual Summer School hosted by the National Association of Extension Program and Staff Development Professionals and the Extension Southern Region Program Leadership Network. Journal articles have been published and are cited throughout this paper. Additionally, the environmental scan survey was shared with other groups interested in using the process in other health-related areas such as chronic disease management programs. Finally, we collaborated with Extension colleagues at the University of Tennessee on a Rural Health and Safety Education grant to implement one of the programs identified through this SSA project. Such work will continue as opportunities arise.

Conclusions

The project reported here provided a snapshot of 4-H Healthy Living programs across the nation at one point in time and led to the identification of 32 promising health-related positive

youth development programs. The review process allowed for the identification of programmatic weaknesses and resulted in recommendations for program improvements. Weaknesses identified in programs provide direction for professional development. This inventory is primarily useful to 4-H as the need for health-related youth programs increases and the pressure to implement programs that have demonstrated positive health outcomes mounts. However, these programs could be replicated by any positive youth development organization. Additionally, other organizations could use the SSA method to conduct an environmental scan and evaluability assessment of their own program offerings. The promising 32 programs have the potential to foster a healthy lifestyle that influences immediate and long-term health outcomes.

Acknowledgement

References

Arnold, M. E. (2015). Connecting the dots: Improving Extension program planning with program umbrella models. *Journal of Human Sciences and Extension, 3*(2), 48–67. Retrieved from https://docs.wixstatic.com/ugd/c8fe6e_8dc5ad8ded104e858ee022fb1ff055d8.pdf

Corporation for National and Community Service. (2012). *Overview of funding opportunity*. Retrieved from http://www.nationalservice.gov/pdf/12_0210_sif_nofa.pdf

Dawkins, N., Wethington, H., Kettel Khan, L., Grunbaum, J. A., Robin, L., Pitt Barnes, S., . . . Leviton, L. C. (2010). Applying the systematic screening and assessment method to childhood obesity prevention. In L. C. Leviton, L. Kettel Khan, & N. Dawkins (Eds.), The systematic screening and assessment method: Finding innovations worth evaluating. *New Directions for Evaluation, 2010*(125), 33–49. doi:10.1002/ev.319

Dillman, D. A. (2007). *Mail and internet surveys* (2nd ed.). Hoboken, NJ: John Wiley & Sons Inc.

Downey, L. H., Peterson, D. J., Donaldson, J. L., & Hardman, A. (2017). An Extension application of the RE-AIM evaluation framework. *Journal of Extension, 55*(3), Article 3IAW1. Retrieved from https://www.joe.org/joe/2017june/iw1.php

Downey, L. H., Peterson, D. J., Le Menestrel, S., Leatherman, J., & Lang, J. (2015). The systematic screening and assessment method: An introduction and application. *Journal of Extension, 53*(2), Article 2IAW2. Retrieved from https://www.joe.org/joe/2015april/pdf/JOE_v53_2iw2.pdf

Glasgow, R. E., Vogt, T., & Boles, S. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health, 89*(9), 1322–1327. doi:10.2105/AJPH.89.9.1322

Hill, L., McGuire, J., Parker, L., & Sage, R. (2009). 4-H healthy living literature review and recommendations for program planning and evaluation. Retrieved from http://www.4-h.org/resource-library/professional-development-learning/healthy-living-training-guides-resources/

Kettner, P. M., Moroney, R. M., & Martin, L. L. (2017). *Designing and managing programs: An effectiveness-based approach* (5th ed.)*.* Thousand Oaks, CA: Sage.

Knowlton, L. W., & Phillips, C. C. (2013). *The logic model guidebook: Better strategies for great results* (2nd ed.). Thousand Oaks, CA: Sage.

Leviton, L. C., & Gutman, M. A. (2010). Overview and rationale for the systematic screening and assessment method. *New Directions for Evaluation, 2010*(125), 7–31. doi:10.1002/ev.318

National 4-H Council. (2010). *Logic models*. Retrieved from https://4-h.org/professionals/professional-development/logic-models/

National 4-H Council. (2012, December 21). *Request for proposal for 4-H healthy living program scan and evaluability assessment*. Chevy Chase, MD: National 4-H Council.

National 4-H Council. (2016). *Grow true leaders: National 4-H Council annual report 2016*. Retrieved from https://4-h.org/wp-content/uploads/2016/03/2016-Annual-Report.pdf

National 4-H Healthy Living Task Force. (2009). *4-H healthy living strategic framework for program planning and evaluation*. Retrieved from http://www.4-h.org/resource-library/professional-development-learning/healthy-living-training-guides-resources/

Peterson, D. J., Downey, L. H., Leatherman, J., Le Menestrel, S., & Lang, J. (2015). Programmatic environmental scans: A survey based on program planning and evaluation concepts. *Journal of Human Sciences and Extension, 3*(3), 143–150. Retrieved from http://media.wix.com/ugd/c8fe6e_4814bc737da246ea96d1d2067328da91.pdf

Rennekamp, R., & Jacobs, C. (n. d.). *Program design*. Lexington, KY: UK Cooperative Extension Service. Retrieved from https://psd.ca.uky.edu/files/program_design.pdf

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

United Nations Development Fund for Women. (2009). *Guidance notes on carrying out evaluability assessment*. Retrieved from http://erc.undp.org/unwomen/resources/guidance/Guidance%20Note%20-%20Carrying%20out%20an%20Evaluability%20Assessment.pdf

Appendix

Programs with Preliminary, Moderate, or Strong Evidence of Replicability Identified in Figure 1

Program name	References	Category	Level of evidence
4-H ATV Safety http://www.atv-youth.org/	Kroll, M. (n.d.). 4-H ATV safety: Leader's guide. Retrieved from http://www.atv-youth.org	Injury Prevention	Preliminary
4-H Bloco Drum and Dance http://anrcatalog.ucanr.edu/pdf/8427.pdf	Conklin-Ginop, E. L. (n.d.). *Developing afterschool programs for teens: 4-H bloco drum and dance*. University of California Cooperative Extension, Sonoma County. Conklin-Ginop, E. L., Braverman, M. T., Caruso, R., & Bone, D. (2011). Bringing Carnaval drum and dance traditions into 4-H programming for Latino youth. *Journal of Extension*, 49(4), Article 4IAW1. Retrieved from http://www.joe.org/joe/2011august/iw1.php Conklin-Ginop, E. L., Junge, S. K., & Pulley, K. (2012). *Part 1: Introduction to 4-H youth bloco drum and dance* [Publication 8427]. Retrieved from http://anrcatalog.ucanr.edu/pdf/8427.pdf	Social-Emotional Health and Well-Being	Preliminary
4-H Health Jam https://4h.extension.illinois.edu/programs/health-nutrition/health-jam	Michigan State University 4-H. (2013). *Jump into foods and fitness*. East Lansing, MI: Michigan State University Extension. University of Illinois Extension. (2014). *4-H volunteers: Get up and move*. Retrieved from http://web.extension.illinois.edu/state4h/volunteers/getupandmove.cfm University of Illinois Extension. (2014). *Nutrition & health.* Retrieved from http://web.extension.illinois.edu/state/programarea.cfm? ProgramAreaID=7 University of Illinois Extension. (2007). *Wellness ways.* Urbana, IL: University of Illinois Extension.	Physical Activity and Healthy Eating	Preliminary

Program name	References	Category	Level of evidence
	University of Kentucky Cooperative Extension. (2007). *Get moving Kentucky*! Lexington, KY: University of Kentucky. University of Missouri Extension. (2013). *Show me nutrition*. Retrieved from http://extension.missouri.edu/p/SMN100		
4-H LIFE http://extension.missouri.edu/4hlife/	Day, T., Small, S., & Fitzsimmons, E. (1987). *Family times manual*. Madison, WI: University of Wisconsin Extension. Trotta, R. (2008). *Tackling the tough skills: A curriculum for work and life*. Columbia, MO: University of Missouri Extension. University of Missouri Extension. (n.d.). *Building strong families: Challenges and choices*. Columbia, MO: University of Missouri Extension.	Social-Emotional Health and Well-Being	Preliminary
4-H Tech Wizards	Oregon State University (OSU) Extension. (2015). *4-H tech wizards replication*. Retrieved from http://extension.oregonstate.edu/washington/4h/4-h-tech-wizards-replication Oregon State University (OSU) Extension. (2015). *Core principles of the program*. Retrieved from http://extension.oregonstate.edu/washington/4h/core-principles-0 Oregon State University (OSU) Extension. (2015). *Washington county 4-H tech wizards*. Retrieved from http://extension.oregonstate.edu/washington/4h//techwizards Oregon State University (OSU) Extension. (n.d.). *4-H tech wizards logic model*. Retrieved from http://extension.oregonstate.edu/washington/4h/sites/default/files/twreplicationmanual.2.pdf Oregon State University (OSU) Extension. (n.d.). *Program goals*. Retrieved from http://extension.oregonstate.edu/washington/4h/sites/default/files/twreplicationmanual.3.pdf	Social-Emotional Health and Well-Being	Preliminary

Program name	References	Category	Level of evidence
4-H Thrive http://4h.ucanr.edu/About/Framework/PYD/Thrive/	Miner, G., & Dogan, S. (2011). *iThrive*. Davis, CA: University of California Division of Agriculture and Natural Resources, 4-H Youth Development Program. Miner, G., & Dogan, S. (2012). *iThrive 2*. Davis, CA: University of California Division of Agriculture and Natural Resources, 4-H Youth Development Program. Miner, G., & Iaccopucci, A. (2014). *iThrive 4.* Davis, CA: University of California Division of Agriculture and Natural Resources, 4-H Youth Development Program. Worker, S., & Miner, G. (2013). *iThrive 3*. Davis, CA: University of California Division of Agriculture and Natural Resources, 4-H Youth Development Program.	Social-Emotional Health and Well-Being	Preliminary
4-H Youth and Families with Promise https://utah4h.org/files/Projects/specialprograms/YFP/ProgramGuide.pdf	Dart, P. C. (2006). *4-H mentoring: Youth and families with promise program guide* (2nd ed.). Logan, UT: Utah State University Extension. Retrieved from https://utah4h.org/files/Projects/specialprograms/YFP/ProgramGuide.pdf	Social-Emotional Health and Well-Being	Moderate
Calcium, It's Not Just Milk https://www.unce.unr.edu/publications/files/hn/2014/cm1401.pdf	Wilson, M. (2001). *Calcium, it's not just milk!* Reno, NV: University of Nevada Cooperative Extension.	Healthy Eating	Preliminary
Choose Health Action Teens	Crosiar, S., & Wolfe, W. (2013). *Choose health action teens*. Ithaca, NY: Division of Nutritional Sciences, Cornell University. Retrieved from	Physical Activity and	Preliminary

Program name	References	Category	Level of evidence
https://fnec.cornell.edu/for-partners/programs/chat/	https://fnec.cornell.edu/Our_Initiatives/CHAT.cfm Crosiar, S., & Wolfe, W. (2013). *Choose health: Food, fun and fitness*. Ithaca, NY: Division of Nutritional Sciences, Cornell University. Retrieved from https://fnec.cornell.edu/Our_Initiatives/CHFFF.cfm	Healthy Eating	
Choose Health: Food, Fun, and Fitness http://fnec.cornell.edu/for-partners/curricula/chfff/	Crosiar, S., & Wolfe, W. (2013). *Choose health action teens*. Ithaca, NY: Division of Nutritional Sciences, Cornell University. Retrieved from https://fnec.cornell.edu/Our_Initiatives/CHAT.cfm Crosiar, S., & Wolfe, W. (2014). *Choose health: Food, fun, & fitness. Healthy eating and active living for 8-12 year-olds: Six hands-on, interactive lessons with food and games*. Ithaca, NY: Division of Nutritional Sciences, Cornell University. Retrieved from https://fnec.cornell.edu/Documents/CHFFF/CHFFF_Intro_FINAL.pdf Wolfe, W. (2015). *Practice-based evidence for Cornell's Choose health: Food, fun, and fitness (CHFFF) curriculum*. Ithaca, NY: Division of Nutritional Sciences, Cornell University.	Physical Activity and Healthy Eating	Preliminary
DARE to be You: Care to Wait http://dtby.colostate.edu/Care%20to%20Wait/Care-to-Wait.shtml	Colorado State University Cooperative Extension. (n.d.). *Care to wait program for families with middle school youth*. Cortez, CO: DARE to be You Program.	Social-Emotional Health and Well-Being	Preliminary
EatFit http://www.eatfit.net/	Horowitz, M., Shilts, M., & Townsend, M. (2009). *EatFit*. Davis, CA: University of California Agricultural and Natural Resources.	Physical Activity and Healthy Eating	Preliminary

Program name	References	Category	Level of evidence
Energy Express https://extension. wvu.edu/youth-family/youth-education/energy-express	West Virginia University (WVU) Extension. (2015). *Energy express: Children*. Retrieved from http://energyexpress.ext.wvu.edu/get-involved/children West Virginia University (WVU) Extension. (2015). *Energy express: Purpose*. Retrieved from http://energyexpress.ext.wvu.edu/purpose West Virginia University (WVU) Extension. (2015). *Energy express: What is energy express?* Retrieved from http://energyexpress.ext.wvu.edu/	Social-Emotional Health and Well-Being	Preliminary
Family Fitness Program	The Pennsylvania State University. (2010). *Family fitness*. State College, PA: Pennsylvania State University Extension. U.S. Department of Agriculture. (2010). *USDA's MyPlate*. Washington, DC: U.S. Department of Agriculture. Retrieved from http://www.choosemyplate.gov	Physical Activity and Healthy Eating	Preliminary
Fast Track	Eunice Kennedy Shriver National Institute of Child Health and Human Development. (2005). *Media-smart youth: Eat, think, and be active!* Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Institute of Child Health and Human Development. Hopper, C., Fisher, B., & Munoz, K. D. (2008). *Physical activity and nutrition for health*. Champaign, IL: Human Kinetics Publishers. Klotzbach-Shimomura, K., & Keenan, D. P. (2000). *Jump start your bones: A school-based osteoporosis prevention curriculum designed for middle school students.* New Brunswick, NJ: Rutgers Cooperative Extension and New Jersey Agricultural Experiment Station. Purdue University. (2002). *Fantastic foods*. West Lafayette, IN: Purdue University.	Physical Activity and Healthy Eating	Preliminary
Food and Nutrition	Gilmer County Cooperative Extension. (2005). *Healthy lifestyles.* Athens, GA: The University of Georgia College of Agricultural and Environmental Sciences.	Physical Activity and	Preliminary

Program name	References	Category	Level of evidence
Education for Children		Healthy Eating	
Gearing Up for Safety http://www.ydae.purdue.edu/tractor/default.htm	Fetzer, L. M. (2013). *Gearing up for safety*. Retrieved from http://www.extension.org/pages/67301/gearing-up-for-safety#.VfLo7xFVhHy Purdue University College of Agriculture. (2008). *Gearing Up For Safety: Production agriculture safety training for youth*. West Lafayette, IN: Purdue Research Foundation.	Injury Prevention	Preliminary
Get Moving–Get Healthy with New Jersey 4-H http://nj4h.rutgers.edu/getmoving/	Devitt, A., Gore, S., Mansue, K., Krzyzanowski, V., & Makres, S. (2005). *Get moving–get healthy with NJ 4-H*. Rutgers, NJ: Rutgers University Agricultural Experiment Station.	Physical Activity and Healthy Eating	Preliminary
Health Rocks! https://4-h.org/parents/healthy-living/health-rocks/	National 4-H Council. (2014b). *Healthy living: Health rocks!* Retrieved from http://www.4-h.org/youth-development-programs/kids-health/programming-resources/preventative-health-safety/health-rocks/	Alcohol, Tobacco, and Other Drug Prevention	Strong
Healthy Weights for Healthy Kids https://ext.vt.edu/food-health/family-nutrition-	Serrano, E., & Jamison, K. (2011). *Healthy weights for healthy kids*. Blacksburg, VA: Virginia Cooperative Extension.	Physical Activity and Healthy Eating	Preliminary

Program name	References	Category	Level of evidence
program.html (Programs tab)			
Jump into Foods and Fitness https://www.canr.msu.edu/jump_into_foods_and_fitness_jiff/	Baird, K., Branta, C., Mark, C., & Seremba, D. (2003). *Jump into foods and fitness*. East Lansing, MI: Michigan State University Extension, 4-H Youth Development, Family & Consumer Sciences.	Physical Activity and Healthy Eating	Preliminary
Just Be It! Healthy and Fit http://babysfirstwish.aces.nmsu.edu/resources.html	New Mexico State University. (2006). *Just be it! Healthy & fit program curriculum*. Retrieved from http://aces.nmsu.edu/ces/nmcyfar/resources.html	Physical Activity and Healthy Eating	Preliminary
National Safe Tractor and Machinery Operation https://extension.psu.edu/national-safe-tractor-and-machinery-operation-program	The Pennsylvania State University, The Ohio State University, & the National Safety Council. (2013). *National safe tractor and machinery operation program* (2nd ed.). Apple Valley, MN: Hobar.	Injury Prevention	Preliminary
Nutrition to	Morris, J., & Zidenberg-Cherr, S. (2013). *Nutrition to grow on: A garden-enhanced*	Healthy Eating	Moderate

Informing Youth Health and Well-Being Programs

Program name	References	Category	Level of evidence
Grow On https://www.cde.ca.gov/ls/nu/he/nrttogrow.asp	*nutrition education curriculum for upper elementary school children*. Sacramento, CA: California Department of Education. Retrieved from http://www.cde.ca.gov/ls/nu/he/nrttogrow.asp		
On the Move to Better Health file:///C:/Users/dp846/Downloads/18state-jgr-ndfh-otm.pdf	Garden-Robinson, J. (2007). *On the move to better health*. Fargo, ND: North Dakota State University Extension Service.	Physical Activity and Healthy Eating	Preliminary
Project ALERT https://www.projectalert.com/	BEST Foundation. (2013). *Project ALERT*. Retrieved from http://www.projectalert.com/	Alcohol, Tobacco, and Other Drug Prevention	Strong
PROSPER https://extension.psu.edu/prosper-strengthens-families	BEST Foundation. (2013). *Project ALERT*. Retrieved from http://www.projectalert.com/ Channing Bete Company, Inc. (2014). *PATHS (Promoting alternate thinking strategies) program*. Retrieved from http://www.channing-bete.com/prevention-programs/paths/paths.html Hockaday, C. (2008). *Strengthening families program: For parents and youth 10-14*. Ames, IA: Iowa State University Extension. Lions Clubs International Foundation. (2014). *Lions quest*. Retrieved from http://www.lions-quest.org/	Alcohol, Tobacco, and Other Drug Prevention	Strong

Program name	References	Category	Level of evidence
	National Health Promotion Associates (n.d.). *Botvin life skills training*. Retrieved from http://www.lifeskillstraining.com/		
Smart Bodies http://www.lsuagcenter.com/topics/food_health/education_resources/smart_bodies	Blue Cross and Blue Shield of Louisiana & LSU AgCenter. (2014). *Louisiana body walk*. Retrieved from http://www.smartbodies.org/smart-bodies/louisiana-bodywalk International Life Sciences Institute Research Foundation. (2012). *Take 10!* Retrieved from http://www.take10.net/ The OrganWise Guys, Inc. (2014). *The organwise guys*. Retrieved from http://organwiseguys.com/	Physical Activity and Healthy Eating	Strong
Strengthening Families https://www.extension.iastate.edu/sfp10-14/	Hockaday, C. (2008). *Strengthening families program: For parents and youth 10-14*. Ames, IA: Iowa State University Extension.	Alcohol, Tobacco, and Other Drug Prevention	Strong
Teen Cuisine https://ext.vt.edu/food-health/family-nutrition-program.html (Programs tab)	Carrington, A. C., & Margheim, L. (2013). *Teen cuisine*. Blacksburg, VA: Virginia Cooperative Extension.	Physical Activity and Healthy Eating	Preliminary
Teen Interactive Theatre Education	University of Arizona Maricopa County Cooperative Extension. (n.d.). *Teen interactive theater education (TITE curriculum)*. Tucson, AZ: University of Arizona Cooperative Extension.	Alcohol, Tobacco, and Other Drug	Preliminary

Program name	References	Category	Level of evidence
(TITE) https://joe.org/joe /2011february/iw5 .php	Wyman Center. (n.d.). *Teen outreach program*. Eureka, MO: Wyman Center, Inc.	Prevention	
Youth Teaching Youth http://www.extens ion.umn.edu/yout h/mn4-H/youth-teaching-youth	Duncomb, M. E. (2010). *Youth teaching youth: Alcohol & tobacco decisions*. 4-H National Headquarters Program of Distinction manuscript, 1–13. Landrieu, J. (2012). *Minnesota 4-H youth teaching youth (YTY)*. St. Paul, MN: University of Minnesota Extension. Retrieved from http://www.extension.umn.edu/youth/mn4-H/youth-teaching-youth/docs/youth-teaching-youth-teen-teachers-report-full.pdf	Alcohol, Tobacco, and Other Drug Prevention	Preliminary