USING THE COMMON MEASURES EVALUATION TOOL IN NORTH DAKOTA: A QUALITATIVE STUDY

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Amelia Kaye Doll

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Program:
Extension Education

October 2019

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Using the Common Measures Evaluation Tool in North Dakota: A Qualitative Study

By

Amelia Kaye Doll

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Adam Marx
Chair

Erika Beseler Thompson

Meagan Scott

Claudette Peterson

Approved:

10/7/19
Date

Chris Ray
Department Chair

ABSTRACT

Using a basic qualitative approach, interviews were conducted with five agents in North Dakota with a range of experiences and lengths of service to evaluate the effectiveness and implementation of Common Measures, a set of survey instruments designed to assess the impact of the 4-H program. The study addresses the culture of North Dakota Extension's views on evaluation and the implementation of Common Measures. The use of a state-wide reporting tool to assess the 4-H program in North Dakota was perceived as necessary, however, many agents felt Common Measures, missed the mark and did not meet the North Dakota reporting. The ability to tell the story of the 4-H program and the long-term impact it has on youth is necessary to the success of any state-wide reporting venture. Additional training on how alternative forms of data collection can be used to tell the story of their program is needed.

ACKNOWLEDGEMENTS

It is often said that no person is an island, completing my master's degree has made that clearer than ever before. It would be hard to list everyone who has helped me along the journey, however I will attempt to convey my thank you to those who have pushed me (sometimes kicking and screaming) to where I am today.

To my participants, thank you for being willing to chat with me about Common Measures. Each agent I spoke with along this journey was open and honest, making my data even more rich. Thank you for sharing your knowledge with me as we continue to make the best better.

My committee; Dr. Adam Marx, Dr. Erika Beseler Thompson, Dr. Meagan Scott, and Dr. Claudette Peterson, your guidance along this journey has been amazing, no words can express what it means to me to have this team of esteemed people behind you, cheering you on. Adam, thank you for guiding me along the path that neither one of us were really sure about in the beginning. Erika's class was the first place I fell in love with qualitative research. To say it was love at first sight, would be a huge lie, however through her encouragement and dedication to helping all students succeed, my love for the process formed. Even when a week before a final project was due and I completely changed my final project, Erika made time to help me navigate the rough waters and come out better on the other side. As a fellow boy mom and qualitative nerd, Erika showed me that this could be done. Meagan was the reason I made said change mid-stream. From the late-night chats as roommates at conferences, to the text, phone calls, and skype meetings, she was always there encouraging me and pushing me. Having a friend with such faith and dedication in seeing you succeed is such an amazing feeling. Claudette was one of my first professors along this journey, from our talks about evaluation in class, to the discussion about the science behind NOT canning in the dishwasher, class was never dull. From the bottom of my heart, thank you all for your support, guidance, and pushing along the way.

None of this would have been possible without my family. Zach, thank you for encouraging me to keep going, for holding down the fort when I needed to retreat away to work on homework, and for believing in me. I know you will not miss the bag of text books and piles of papers in the house. William and Owen, thank you for understanding that mom needed to work on homework. I look forward to more adventures (and board games) with you with far less textbooks along for the ride. Mom, thanks for the encouragement and being there to talk positive youth development with me. Yes, you were right, I should have gone into education from the beginning.

To friends who supported me. Angie, thank you for the encouraging words and support. It never failed if I second guessed myself or felt like it was never going to end, Angie was there to prove me wrong and push me to go farther and do better. Stephanie for the Sunday morning coffee dates that revolved around our homework (and maybe a little chatting), thank you for motivating me to make sure I made the time to get the work done. Bob, I am going to miss our random gif conversations that got us through class some nights.

To all the family and friends that pitched in along the way, we did it. No person is an island, I did not do this alone, and for you I am extremely grateful, thankful, and oh so loved.

DEDICATION

To my kiddos, William and Owen, may you always understand the value of hard work and dedication. As you move forward in life, may no one ever tell you that you can't do something or you are not smart enough for something, those people are wrong. May you always have a passion for learning and following your dreams.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

After the war of 1812, Americans started to question the types of education provided at Universities in the newly formed union. Due to European influence, higher education institutions focused on law, ministry, and teaching the classics, all of which were not abundantly useful in a growing agricultural country. A slow sentiment grew for the need to have a different type of education provided, one that would prepare men for ordinary vocations such as agriculture. The process began slowly, and it was not until 1862 that Senator Morrill was able to successfully pass the Morrill Act, the first of many acts that were needed to formalize what is now known as the Land Grant University system, it started a chain reaction (Sanders, 1966).

New universities faced several up-hill battles due to the education provided. Often with no textbooks, curriculum, or professors, many Land Grant universities struggled in the beginning. Finally, in 1887, congress responded with support for this program through the signing of the Hatch Act. This act, which established an agricultural experiment station that was connected to one university in each state, provided the Land Grant Universities with much needed money that would be used to develop new courses and research advancements in the agriculture field (Sanders, 1966).

The development of the experiment stations led nicely to demonstration farms, a technique to demonstrate new farming techniques to farmers. The farms were first started in 1902 by Seaman A. Knapp, who, at the time was a special agent for the Promotion of Agriculture in the South. Over the next 4 years, Knapp's continued success lead to other counties in Texas offering to pay to have an agent in their county. During the same time, African Americans were hired by the Tuskegee and Hampton institutes to work with African Americans residents in a similar fashion to Knapp and his colleagues (Sanders, 1966).

At the turn of the century, families living in rural communities often only educated their youth in a formal school setting until about eighth grade. After eight years of education most rural boys were expected to help on the farm while the girls became homemakers. A growing number of rural educators felt they were not providing practical and valuable education for the students' futures. To address this concern, community clubs with a focus on "learning by doing" provided opportunities for practical learning that help prepare the youth for the future and at the same time address community concerns. In the beginning, the club leaders were often the school teacher. Meetings focusing on specific projects such as; crops, livestock, canning, sewing, and related project areas. Unlike clubs today, there was little focus on the social activities, group actions, or recreation. An individual approach to project development allowed the youth to address concerns that were tailored to their personal and community needs. In 1909, Knapp recognized the power these clubs had and charged the county agents with club management and supervision (Sanders, 1966).

When the Smith-Lever act was signed in 1914, education in the rural communities continued to evolve. This act would combine all the work Knapp and other agents were doing under one name, the Cooperative Extension System. This system would combine the research and knowledge of the already established Land Grant universities with the skills and delivery of Knapp and others to create a unified Extension system. This provided permanent appropriation to ensure funding was in place to provide the necessary education (Sanders, 1966).

With the development of the Extension system, the boys and girls clubs that Knapp had started working with were also changed. Rural teachers looked to the representatives of the agriculture colleges as resources that could be used in the club meetings. This natural partnership was further solidified with the passage of the Smith-Lever Act in 1914. The passage of this act led to the coordination of the education efforts of the local project-based clubs. These clubs would soon evolve into the boys' and girls' club program of the Extension system, later to be known as 4-H. The youth clubs continued to grow and refine what it meant to be in an Extension club. Over the next fifteen years, emphasis was put on developing the club experience. This meant meetings on a regular basis, elected offices, and a more coordinated approached to educational programing. With a focus on year-round involvement, adult leadership of the club shifted from school teacher to community volunteer, often a farmer, farmer's wife, or parent. This shift allowed the clubs to meet year around and continue to address community needs as they arose.

Areas of study have grown significantly since the early years of Cow Clubs with youth focusing on everything from robotics and science, to baking, sewing, and protecting the environment. Now more than 100 years later, the 4-H program serves youth across the nation in rural, urban, and suburban communities (National 4-H Council, 2019).

With all of the change and progression of the 4-H program, the need to demonstrate the impact the 4-H program has on those participating has remained the same. However, the strategies of program evaluation continues to evolve and grow. Complicating these strategies further is the evolution of complex staffing, local, state, and federal funding sources, program delivery methods, and organizational structures all of which create a more challenging approach to evaluation (Franz & Townson, 2008).

Lerner and Lerner (2013) were the first to address the growing concern of reporting the impact of the 4-H program on a large scale. *The 4-H Study of Positive Youth Development* was a longitudinal study, repeated annually from 2002 to 2008. The study focused on three areas; "positive and sustained relationships between youth and adults, activities that build important life skills, and opportunities for youth to use these skills as participants and leaders in valued community activities" (Lerner & Lerner, 2013, p. ii). Extension Agents point to the Lerner study as the basis for much of the work that is done in 4-H Youth development. Other research has been conducted on youth development, but never to the scale as the Lerner study.

A tool that could be used by state 4-H programs to evaluate local, state, and regional programing efforts was developed in 2012. Over the course of the next six years, the National 4-H council worked on the development and refinement of what is now Common Measures 2.0. The Common Measures surveys provide Extension Agents and state staff with the tools needed to showcase the work that is being done as part of the 4-H Youth development Program (National 4-H Council, 2015). Common Measures currently has 14 instruments that are used in conjunction with a demographics survey (Hawley, 2017).

1.1. Statement of The Problem

Evaluation within the 4-H youth development program poses several challenges and opportunities for Extension Agents. The development and implementation of Common Measures, a set of survey instruments designed to assess the impact of the 4-H program, by the National 4-H Council was meant to minimize the stress that is caused by evaluation of youth and increase fidelity of the 4-H program throughout the nation (National 4-H Council, 2015).

Currently in North Dakota, agents are asked to evaluate programming using the four levels of the Kirkpatrick model which focuses on behavior change within a training program. Originally developed for use in a HR setting, Kirkpatrick's four levels are used in Extension to evaluate and report, the effectiveness of a program (Kirkpatrick, 1996). The problem with this system in relation to 4-H youth development, is that often not one specific program or interaction produces a high-level impact, but rather a series of programs and interactions over an extended period with Extension staff and 4-H volunteers that creates lasting results.

A disconnect is seen between the type of data that is collected using Common Measures and what is required of North Dakota Extension agents for their annual reports. The lack of context and training around evaluation in North Dakota Extension combined with the increased pressure to report solid (Level 3 or 4 on Kirkpatrick's model) impacts leaves agents unsure of how to meet this need. The implementation of Common Measures in North Dakota during the 2017-2018 4-H year will serve as the framework to evaluate the barriers which exist to implementation of an effective evaluation process in 4-H youth development in North Dakota.

North Dakota 4-H finds itself in a unique situation with evaluation. At the forefront of the issue, North Dakota 4-H has a tool, Common Measures, which provides an opportunity for data collection on a large scale state wide. This data has the ability to articulate the lifelong impact the 4-H program has on members across the state. However, this opportunity is overshadowed by the requirement to report impacts using Kirkpatrick's four-level model.

1.2. Purpose of The Study

This research will examine the implementation of the Common Measures assessment tool by North Dakota Extension Agents through a qualitative review of the current use of the tool, the benefits and drawbacks of Common Measures, as well as ways North Dakota Extension and Common Measures can be improved moving forward. Areas for professional development related to evaluation and reporting will assist North Dakota Extension with future state side initiatives.

1.3. Research Questions

Five research questions framed this study:

1. Prior to Common Measures, how did Extension agents evaluate their 4-H program?
2. What are Extension agents' perceptions of the importance of evaluation in their role as an agent?
3. Do Extension agents view new programs and/or procedures, such as Common Measures, as intrinsic or extrinsic motivated?
4. What training or professional development is needed in North Dakota to improve evaluation of the 4-H program?
5. How does an agent's length of service with Extension, position, education level, and county composition influence differing viewpoints of evaluation?

1.4. Conceptual Framework

"To evaluate something means determining its merit, worth, value, or significance" (Patton, 2008a, p. 5). Patton (2008a) identified a large problem with evaluation, is getting people to use the knowledge that is produced. Gathering the data and checking the box when the task is complete is important, however using the evaluation information collected to show significance and improve future programs makes an evaluation tool stronger (Patton, 2008a). Providing Extension professionals with the tools to complete evaluation is one step, providing training and resources to use the tools to make meaningful change in their 4-H program is one step further.

The real magic of evaluation happens when there is a shared definition and commitment to the process (Bailey & Deen, 2002). During a training session on evaluation, Bailey and Deen (2002) found that without a foundation of the basics of evaluation, agents that completed the training did not fully understand how or why the new system being introduced was useful. As the emphasis on evaluation in Extension grows, it is important to also provide staff with the tools necessary to use the information that is collected (Bailey & Deen, 2002).

Within Extension programs, there is a growing need to demonstrate the value of a program in order to obtain or maintain funding (Stup, 2003). Just as Extension agents are competing for funding, they are also competing for time. It is important to make sure any program, new or old, is worth an agent's time. One way to do this is through evaluation. Evaluation is key to seeing what differences a program or event made, if change happened, and how it can change for the future (Hill & Seger, 2018). Patton (2008b) pointed out that Extension has always been in the business of adaption and change, which includes an ever-changing measure of success. Patton (2008b) also pointed out that only when people see evaluation from their world view, do they begin to understand its importance and power.

1.5. Significance of The Study

This study will examine not only how Common Measures can better be utilized by Extension Agents in the field, but also how future evaluation tools can be developed and implemented across the state in an effective and meaningful way. As it stands, the data collected using Common Measures, is not meeting the needs of county agents in terms of required state reporting. While this data is significant for state reports, agents often do not see the connection, and thus have little buy in to the evaluation process. By providing training and education to Extension agents, there is an opportunity for both the state and local programs to benefit from the data collected using the Common Measures tool.

1.6. Definitions of Terms

4-H logic model: A model of youth development that assesses the situation, inputs, activities, and outputs which lead to outcomes in knowledge, actions, and conditions (National 4-H, 2010). The 4-H logic model was used as the basis for the Common Measures instruments.

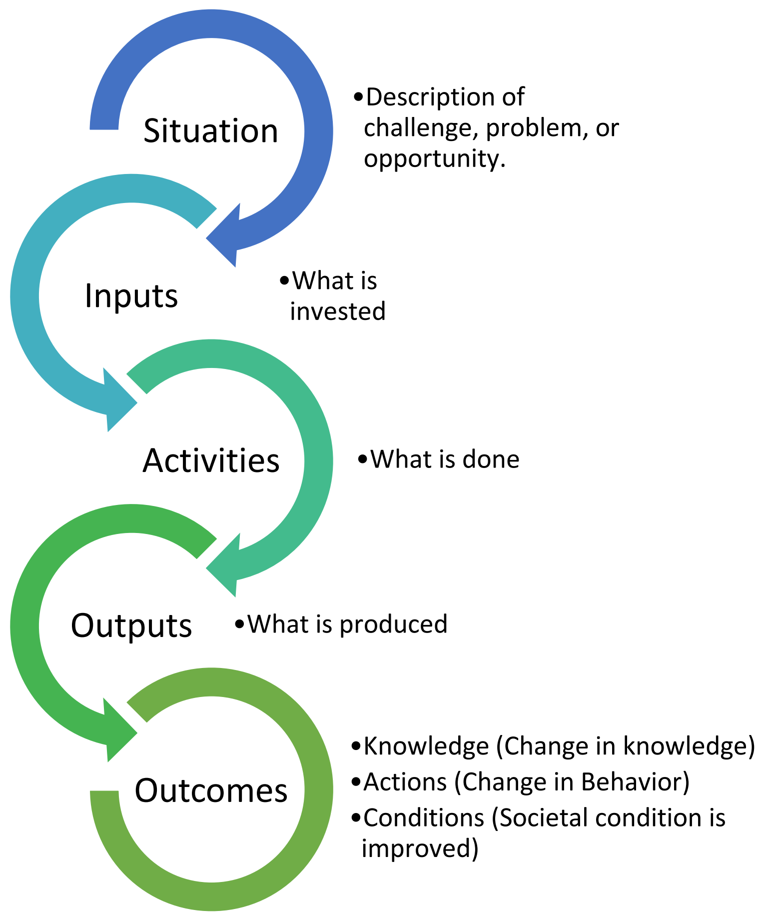


Figure 1. 4-H logic model.

4-H youth development: A youth development program delivered by the Cooperative Extension service. Based on a youth's interest, youth ages 8-18 (5-7 as Cloverbuds) are guided by adult mentors to develop their own learning path (National 4-H Council, 2019). At the center of the 4-H youth development field is what is known as the essential elements (belonging, independence, mastery, and generosity) (Martz, Mincemoyer, & McNeely, 2016).

Assessment: An assessment is a method or tool that is used to evaluate, measure, and document the readiness, learning progress, skill acquisition, or educational needs of learners (Great School Partnership, 2014). Due to the nature of Extension, many agents use the terms assessment and evaluation interchangeably. While there is a distinct difference between the two, in the study, those interviewed had varying definitions of assessment and evaluation. For this reason, you will see both terms used in the results section.

Essential elements: 4-H defines a successful learning experience as one that creates a sense of belonging, independence, mastery, and generosity (Martz, et. all, 2016). Essential Elements was also used as a basis for the Common Measures instrument.

Evaluation: In this study, Fournier's (2005) definition of evaluation will be used.

> An applied inquiry process for collection and synthesizing evidence that culminates in conclusions about the state of affairs, value, merit, worth, significance, or quality of a program, product, person, policy, proposal, or plan. Conclusions made in evaluations encompass both an empirical aspect (that something is the case) and a normative aspect (judgment about the value of something). It is the value feature that distinguishes evaluation from other types of inquiry (p. 139-140).

Extension Agent: Field staff that use knowledge of local issues and citizen input to serve counties across the state (NDSU Extension, 2019).

4-H youth development Extension Agent: An Extension agent with a focus on the 4-H program and positive youth development.

Agriculture natural resource (ANR) Extension Agent: An Extension agent with a focus on the agriculture natural resource program areas.

Family community wellness (FCW) Extension Agent: An Extension agent with a focus in the areas of family and community wellness.

Extension Specialist: With technical expertise in their discipline, specialist develop and lead programs that address statewide concerns (NDSU Extension, 2019).

Common Measures: A set of self-report surveys designed for youth in fourth through twelfth grades, to describe a youth's' 4-H experiences and evaluate the effect of the 4-H programming (Hawley, 2017). The survey can be used with any youth who participated in a 4-H program, project, or grant for at least six participation hours (Hawley, 2017).

Cooperative Extension service: A partnership between the United States Department of Agriculture (USDA) and the National Institute of Food and Agriculture (NIFA) that brings the research from the lab into the communities where people are putting the knowledge into practice. Cunty-based agents empower, educate, and facilitate, farmers, ranchers, youth, and communities to meet the challenges they face (National Institute of Food and Agriculture, 2019a).

Impact statements: Impact statements are used by NDSU Extension to show the economic, environmental, or social outcomes of educational programs. Extension agents in North Dakota are required to submit at least one impact statement per year that reports at least an intent to change behavior or implement new information (NDSU Extension, 2008).

Land grant university: A land grant university is a college or university that has received the designation from the state legislature to receive the benefits of the Morrill Acts. These Universities were originally started to provide education for agriculture and related fields (Sanders, 1966). North Dakota State University (NDSU), formally known as North Dakota Agricultural College, holds this distinction in North Dakota

Narratives: A monthly or quarterly reports required by the North Dakota Century Code that is shared with the county commission, the district director, and in some cases the county advisory board and 4-H council. The report focuses on "program successes and impacts" (NDSU Extension Service, 2017, p. 1).

Packaged program: A developed program that can be delivered by field staff, most often an Extension agent, that addresses a need. Packaged programs often come with all necessary lesson plans and evaluation tools to deliver the program.

2. REVIEW OF LITERATURE

Proper reporting of the impacts made by a program is key to the success of any organization, including Extension. The increasing pressure of Extension personnel to report impacts of their programs creates an even bigger challenge. This chapter provides an in-depth review of the literature related to the study's key variables. In this chapter, there are six main sections addressing: The Land Grant System, Cooperative Extension, what Extension agents do, expectations of Extension agents, and evaluation.

2.1. The Land Grant System

After the War of 1812, it was quickly realized that the people with knowledge and education, who had primarily come from England, were now missing. Jonathan Turner from Illinois was vocal for the need to provide education for the working man, an area not addressed in the current educational system. Up until this point, education was largely reserved for the elite members of society with a primary focus on ministry, medicine, and law. A few small colleges with a focus on agriculture and mechanical instruction started in the mid-1800s. However, it was not until a second attempt at passing the Morrill Act in 1862 that the Land Grant University system was established (Sanders, 1966).

Establishment of the Land Grant System provided rural communities with more opportunities for education than ever seen before. However, with these opportunities came several hurdles to the success of the Land Grant System. Many of the newly established colleges did not have textbooks, courses of study, and most importantly, they were missing professors. They also soon realized that the traditional learning model of lecture was not effective with this type of information (Sanders, 1966).

In North Dakota the Land Grant university is North Dakota State University (NDSU). Originally started in 1890 as North Dakota Agriculture College (NDAC), a name it would continue to operate under until 1960. A year later, in 1891, the first students were admitted to NDAC for short courses. On July 1st, 1914, the North Dakota Extension service was formed as a department of NDAC (NDSU Archives, 2019).

The Land Grant system continued to grow and develop over the next several decades. However, it was quickly realized that the organization could only do so much with what it had. The Hatch Act of 1887 and the second Morrill Act of 1890 helped ease some of the growing pains the Land Grant system was experiencing. The Hatch Act of 1887 provided additional resources for Land Grant universities to conduct research. This act authorized each Land Grant university to establish a research station. The research stations would provide an opportunity for students to learn new practices as well as conduct research. While the amount given to these experiment stations was not much ($15,000), the act provided resources to be able to improve and develop courses that would be taught at the universities (Sanders, 1966). The Morrill Act of 1890 provided funding for 19 historically black universities. The second Morrill Act also forbid racial discrimination in admissions policies of any school receiving federal funds. (National Institute of Food and Agriculture, 2019b; Committee on the Future of the Colleges of Agriculture in the Land Grant University System, 1995; Sanders, 1966).

2.2. The Cooperative Extension Service

The Land Grant university system was a stepping stone toward providing education to the rural areas of America. However, a gap still existed for those who lived in the most rural areas and those who did not attend the universities (Sanders, 1966). The Morrill and Hatch Acts laid the framework for what would later be known as the Cooperative Extension Service. In 1914, the

Smith-Lever Act made the new system of the Cooperative Extension official and established a defined relationship between the Cooperative Extension Service and the United States Department of Agriculture (USDA) (National Institute of Food and Agriculture, 2019a).

While the Cooperative Extension Service has grown and changed a lot since its humble beginnings in 1914, much of the mission of Extension remains the same; Extension provides non-formal education and learning activities to people throughout the country — to farmers and other residents of rural communities as well as to people living in urban areas. It emphasizes taking knowledge gained through research and education and bringing it directly to the people to create positive changes (National Institute of Food and Agriculture, 2019a).

2.3. What Extension Agents Do

Due to the county-based nature of Extension, no two Extension agents do the exact same work. However, the underlying goal of Extension is to provide instruction for county residents (Sanders, 1966). How agents choose to do this education also varies. As Sanders (1966) stated, "learning experiences should be selected as carefully and as specifically as a doctor writes a prescription" (p. 111). Extension agents often use multiple approaches to ensure the highest success rate, often using multiple approaches to reach the same end result (Sanders, 1966). "It takes a combination of at least five different methods to obtain the optimum adoption of a recommended practice" (Sanders, 1966, p. 128). Extension agents utilize their skills, networks, and volunteers to accomplish this mission. Some examples of teaching methods Extension agents use to meet the needs of their clients include:

- Client inquiries – Client inquiries most often come in the form of phone calls, walk-ins, or email. Work is often done in a one-on-one setting to address a specific client's needs.

- Field days – Field days are most often used within the agriculture community to provide hands on learning in the field. New cover-crops or no till fields are two focus areas that might be addressed on field days (Sanders, 1966).
- Demonstrations – Demonstrations are the foundation of Extension work in the United States. Demonstrations allow the agent to visually and verbally explain a process, fact, or idea (Sanders, 1966).
- Educational classes – Educational classes range widely in content, audience age, and delivery method. They can be a few hours in length, or span over several weeks or months. Classes are generally small groups that work together on a specific problem or task with an Extension agent, the expert, leading (Sanders, 1966).
- Display booths – Booths at a variety of community events and activities allow the Extension service to disseminate information to the public.
- Publications – Often a few pages long, publications are a tool used to educate or reinforce education provided by Extension agents (Sanders, 1966). Publications are often available in print and electronic form.
- Facilitation – Groups large and small are often in search of someone to facilitate discussions and provide personal or professional development opportunities for their groups.
- School enrichment – Often one of the best places to reach youth is by going to where they are, school. Providing direct education to school youth enriches the classroom experience and is a great opportunity for Extension agents to reach the younger population.

- Certifications – Extension agents are trained to certify clients through the delivery of specific programs or classes. Examples of certifications provided by the Extension service include; pesticide, ServeSafe, and master gardener.
- Contests and fairs – Contests and fairs are used by Extension agents on the local, district, and state levels to showcase the work of members of the community, most often, 4-H youth. Contests and fairs also provide agents with an opportunity to help tell the Extension story (Sanders, 1966).
- Community development – Since 1958, community development has been recognized as an area of program emphasis by the Extension Service (Sanders, 1966). Community development work looks different depending on the needs of the community, but it can include; civic engagement, rural community development, or leadership development (Flage, 2019).

2.4. Expectations of Extension Agents

When a new agent starts with NDSU Extension, there are many expectations that are outlined for them, the primary categories being: educate, collaborate, and facilitate. In partial fulfillment of the educate category, agents are expected to "teach, deliver, and evaluate programs" in their specialty area (NDSU Extension, 2019, para 4). The job description provides additional direction for agents saying they will need to "evaluate and market the impact of the educational programs delivered" (NDSU Extension, 2019, para 8). To allow new agents enough time to develop relationships and programs within their community, new agents are not required to submit an impact statement during their first year of employment (NDSU Extension, 2008).

2.5. Evaluation

A distinguishing characteristic of evaluation is the ability to show value. This characteristic sets evaluation apart from other forms of inquiry such as journalism, basic science research, or polling (Fournier, 2005). Evaluation is defined as

> an applied inquiry process for collecting and synthesizing evidence that culminates in conclusions about the state of affairs, value, merit, worth, significance, or quality of a program, product, person, policy, proposal, or plan. Conclusions made in evaluations encompass both an empirical aspect (that something is the case) and a normative aspect (judgement about the value of something) (Fournier, 2005, p. 139).

Showing value is imperative to continued support of Extension and the programs that are offered throughout the state.

2.5.1. History of Evaluation

Evaluating the success of teaching programs is not a new concept and can be traced back almost as far as formal education itself. Evaluation of teaching programs can be documented more than 4,000 years ago with formal proficiency testing conducted by the emperor of China. While delivery methods have changed and evolved over the years, the objective of evaluation remains the same, to assess something that was done and see how it can be made better (Patton, 2015).

2.5.2. Importance of Evaluation

At its core, evaluating a program can help improve the program and show results to stakeholders. Within the Extension system, there are many stakeholders that agents are responsible for reporting programmatic impacts. For example, a county agent is responsible for reporting impacts to the county as well as to the state. Additionally, an agent could have other partners or funders that require program reports. The need to provide different reports to each stakeholder creates a push and pull between doing impactful work and reporting impacts (Patton, 2008b). These partnerships have become more complex over time, increasing the importance of program evaluation and reporting (Franz & Townson, 2008).

The autonomous nature of Extension adds additional complexity to program development and evaluation (Franz & Townson, 2008). The many working pieces of an Extension program are again complicated by finding the balance between evaluating to prove, and evaluation to improve (Patton, 2008b).

2.5.3. Evaluation Across the Extension System

Across the Extension system, there are a variety of different approaches and thoughts on evaluation and the use of data. A report released by Kushner, Crave, and Jones (2017) reported that even when the demand for "evidence-based decision-making and accountability" increases, many states decrease their professional development commitment (p.8). Decreased budgets have a large role in organizations replacing human capacity with tools and materials (Kushner et. al, 2017).

Many factors go into deciding the methodological rigor of an evaluation tool. Striking a balance between the costs (financial, time, and participant burden) and the benefits (funding requirements, opportunity, standards and trends in the field, program requirements for recognition and legitimation, evaluator expertise, and organizational culture) allows evaluators to find a sweet spot for their evaluation tool (Braverman & Arnold, 2008).

Further complicating a unified approach to evaluation across the country is the need for Extension agents to involve community stakeholders in programmatic development. These collaborations increase the depth and breadth of the impact of local programming (Franz &

Townson, 2008). This again brings more people to the table, creating additional requirements of what is needed for a successful evaluation (Franz & Townson, 2008).

The Land Grant Universities were started to primarily educate and research the agricultural community and the best way to do this was to place university agents into the communities they were serving. Since its beginning, Extension has shown programming impacts on a local level. In fact, a movement was made toward using public value to articulate the impact of Extension programs on the greater good (Franz, Arnold, & Baughman, 2014). Public value shows the impact a program has on more than just those who directly participated in a program (Lewis, Horrillo, Widaman, Worker, & Trzesniewski, 2015). Showing a public impact of a program provides stakeholders with more robust data to justify supporting the programming that is taking place. In North Dakota, the change to add public value statements as a requirement for impact statements was made in 2016.

2.5.4. Evaluation in Extension Programs

Evaluation plays a vital role in Extension programming funding and continuation. Even with the large demand, the connection between evaluation and program planning is not very strong (Kushner, et al., 2017). Within the positive youth development program area, the Lerner study, showed the long-term impact of the 4-H program on youth who participated in the study. While the Lerner study was the first of its kind and it helped the 4-H program in many ways, the data is starting to age (Lerner & Lerner, 2013).

2.5.4.1. Lerner Study

The last multi-state assessment of the 4-H youth development program was completed by Tufts University between 2002-2010. The study, which was completed annually with over 7,000 youth from diverse backgrounds in 42 states, gave researchers the first real look into positive youth development programming (Lerner & Lerner, 2013). Lerner and Lerner (2013) hypothesized that youth who possessed the five Cs of youth development (competence, confidence, character, connection, and caring) were on a developmental path to the sixth c, contribution.

As the first study of its kind, the Lerner study paved the way for research in positive youth development. By providing opportunities for youth to have social nutrients, all young people have the opportunity to work toward a life of successful contributions. Youth development work, specifically 4-H in this case, helps launch "young people into healthy and productive lives" (Lerner & Lerner, 2013, p. 15).

While the Lerner study had some key findings that contribute greatly to the positive youth development field, much has changed for youth in the ten years since the completion of data collection. Given the longitudinal nature of the study, youth from marginally at risk/in risk backgrounds become difficult to reach (Lerner & Lerner, 2013). As the Lerner study ages, it became apparent it was time for a new multi-state assessment of the 4-H youth development program.

2.5.4.2. Common Measures

The Common Measures library of assessments was created to help the 4-H program evaluate the impact that 4-H programming is having on the common core of child/youth outcomes and indicators outlined by the National Institute of Food and Agriculture (NIFA) plan of work. Over five years (2012-2017) collaborators from 4-H National Headquarters, National 4-H Council, and representatives from Land Grant Universities, used the National 4-H Logic Model to develop what is now known as Common Measures 2.0. Revisions of Common Measures 1.0 to Common Measures 2.0 started in the spring of 2017 guided by a team of land grant university faculty representing the national Extension regions (National 4-H Council, 2015). Common Measures, as it stands, “is designed for youth in fourth through twelfth grades who have participated in a 4-H program, project, or grant” for a minimum of six contact hours (Hawley, 2017, pp. 2-3).

In addition to a basic demographics survey, Common Measures 2.0 is broken down into six different instruments, each with a specific target age:

Table 1

Common Measures Instruments

Tool	Number of Questions			Targeted Age
	Core	Supplemental	Total	
Demographic	5	8	13	Grades 4-12
4-H experience	16	--	16	Grades 4-12
Universal	23	--	23	Grades 4-12
Citizenship				Grades 4-12
Interest in Community Service and Giving back	5	9	14	
Community Awareness	5	4	9	
College and career				Grades 8-12
Professionalism	6	8	14	
College Decision Making	6	4	10	
Career Decision Making	10	5	15	
Healthy Living				Grades 4-12
Healthy Eating Habits	6	4	10	
Being Active	4	1	5	
Healthy Decision Making	2	6	8	
Food Preparation	--	6	6	
Science				
Science Interest and Thinking	5	9	16	Grades 4-12
Science Skills and Attitudes	7	--	7	Grades 8-12
Engineering Skills and Abilities	7	--	7	Grades 8-12

(National 4-H Council, 2015)

When an agent or specialist uses Common Measures, all the core questions in a category must be used if the data will be submitted to the state or nation. Supplemental questions can be chosen as needed to meet the specific needs of the local program. Table 2 provides an example of the options an agent looking for an engineering and science focused survey, with demographics would have. The survey could range in length of ten questions (demographics and science interest and thinking with no supplemental questions) to thirty-six questions (demographics, science interest and thinking, and engineering skills and abilities with all supplemental questions).

Table 2

Engineering Program Example

Tool	Number of Questions		
	Core	Supplemental	Total
Demographic	5	8	13
Science			
Science Interest and Thinking	5	9	16
Engineering Skills and Abilities	7	--	7
Totals	17	17	36

Common Measures is used to evaluate the impact of the 4-H program using the Essential Elements framework. This framework addresses the four necessary attributes of youth programs striving to create environments that are conducive to positive youth development.

Belonging Positive Relationship with a caring adult An inclusive environment A safe environment	**Mastery** Engagement in learning Opportunity for mastery
Independence Opportunity to see oneself as an active participant in the future Opportunity for self-determination	**Generosity** Opportunity to value and practice service for others

(Martz, et. all, 2016)

Figure 2. The eight elements of positive youth development distilled to four concepts.

Agents can use the Common Measure tool in a variety of ways including program specific evaluation, as a needs assessment, or as a cross-sectional tool (National 4-H Council, 2018). However, if the tool is going to be used, it is imperative that the instruments not be altered in any way and that each survey be used in its entirety (Hawley, 2017). Consistency across uses allows the data to be aggregated and used for state or nation-wide reporting.

One key factor of using Common Measures is the amount of contact time youth have had with the 4-H program. It is recommended that youth have a minimum of six contact hours before they are administered any of the instruments. This may mean that the instrument is not used for many youth development programs (Hawley, 2017).

The dosage of time spent in programming is key to developing positive youth development skills within youth. “Children and youth will not benefit unless they attend programs regularly, and evidence is emerging that those who attend more frequently and for longer periods of time benefit more than their peers whose attendance is more sporadic” (Fiester, Simpkins, & Bouffard, 2005, p. 91). Not only does a high dosage of youth programs impact the development of positive youth development skills, research also shows that a high dosage has a positive impact on their academic and social behaviors versus youth who participate at a lower dosage (Fiester, et al., 2005).

To assess the validity of cross state programming, programming happening in more than one state with a common goal, a group of seven programs from across the United States came together in 2012 to find a way to evaluate the outcomes of programs funded by Children, Youth, and Families At Risk (CYFAR) grants. This pilot study allowed the group to look at a common evaluation tool being used on programs that vary by content and delivery method but have the same overall goals. To accommodate the variety in the delivery, two formats of evaluation were provided for programs to use; pre/post survey and retrospective (Payne & McDonald, 2012).

This group of Extension professionals developed and delivered an evaluation tool across four of the states. Even if the delivery method varied across the states, the programming still contributed positively to the outcomes evaluated. They found that the overall time participating in the program, which varied by site, had an impact on the outcomes. Youth who participated for at least ten months showed the greatest increase compared to the medium (four to nine months) and low (three or fewer months) groups (Payne & McDonald, 2012).

A further study was done in 2015 by the state of California using the Common Measures tool with 721 youth from across the state through embedding surveys in the state's online record book program. While California was successful in data collection, the group made several recommendations to improve the viability of the tool, many of which were addressed in the Common Measures 2.0 tool. Some recommendations made include, (Lewis, Horrillo, Widaman, Worker, & Trzesniewski, 2015).

Within the North Dakota 4-H program, Extension agents from across the state were tasked with administering the Common Measures instrument to a minimum of 30% of the enrolled 4-H youth in their county. The enrollment in North Dakota 4-H as of August 2018 was 4,810 youth, making the target of 30% of youth surveyed, 1,443. At the completion of the project 2034 surveys were collected across the state. The initial wave of surveys used the universal instrument. Youth completed the survey at a variety of places including county fairs, awards nights, and club meetings. In addition to county-based programming, state specialists also used it for statewide programs and events.

2.5.4.3. Current Evaluation Tools Used by North Dakota State University Extension

Currently, NDSU Extension uses the Kirkpatrick model to determine impacts of programing. The state requires each agent to evaluate at least one program annually (NDSU, 2008a). The Kirkpatrick model categorizes evaluation into four levels of impact; reaction, learning, behavior, and results (Kirkpatrick Partners, 2018). NDSU has added a fifth level, level 2.5, intent, an additional level an agent or specialist can also use to show the impact of their program (NDSU Extension, 2008). Using this model, agents develop tools that are used to show the impact of their programs. Results of these evaluations are placed in one of the four areas as defined by Kirkpatrick:

Level 1 - Reaction: how much did the person like the program. A level 1 does not measure any learning that occurred or behavior change as a result of the training or programming. Questions often asked at this level include: did the time of the program work for you, did the program meet your expectations, was it easy to find the location (Kirkpatrick, 1996).

Level 2 - Learning: “A measure of the knowledge acquired, skills improved, or attitudes changed due to the training” (Kirkpatrick, 1996, p. 56). This level is often described as a knowledge check. Similar to a test in school, this level is used to ensure those in attendance can provide answers to knowledge-based questions (Kirkpatrick, 1996).

Level 2.5 – Intent: a level created by NDSU Extension to show impact that is more than learning but has not quite reached behavior change. Evaluations at this level collect data about participants intent to change behavior given the information received in the program or event (NDSU Extension, 2008).

Level 3 - Behavior: As Kirkpatrick (1996) explained, there is a big difference between knowing something and applying it. Level 3 evaluates how much of the knowledge gained in the training is being used in application. For best results, this data is usually collected three to six months after the education has taken place (Kirkpatrick, 1996).

Level 4 - Results: the final level in Kirkpatrick's model focuses on measurable changes made as a result of the training. Examples can include; high productivity, money saved, or improved quality. Again, this data is often collected several months after the program has taken place (Kirkpatrick, 1996).

As evaluations move from level one to four, the impacts of a program become more impactful and meaningful as well as more expensive, complicated, and time consuming (Kirkpatrick, 1996). As a requirement for reporting agents are asked to submit an impact statement that shows a level 2.5 impact or higher annually (NDSU Extension, 2008). According to the NDSU Extension website, this will change to a level 3 impact for the 2019 reporting year.

In 2016, public values statements were added to the NDSU impact statement format. Having public values statements assists agents and administrators in articulating the broader impact of programs (Franz, et al., 2014). A program has an even broader reach if you also report the social, economic, or environmental condition impact of the larger community.

In addition to the Kirkpatrick model, North Dakota agents are encouraged to use the model (Table 3) created by Merril Eweret, a past director of Cornell Cooperative Extension. In this model, Eweret examines the process as well as the content of a program to assess where agents are spending most of their time (Franz & Townson, 2008).

		Content	
		Low	High
Process	High	**FACILITATION** Community issues, forums, strategic planning (Process evaluation)	**TRANSFORMATIVE EDUCATION** Workforce development, youth obesity, (Impact evaluation)
	Low	**SERVICE** Partnerships, fee for service (Customer satisfaction Evaluation)	**CONTENT TRANSMISSION** Web sites, newsletters, PowerPoint Presentations (Knowledge Evaluation)

*Figure 3.*Transformational education model.

The facilitation area of this model describes when agents are serving as facilitators of group processes in an educational setting rather than serving as the person delivering the program (Franz & Townson, 2008). For example, facilitating council and committee meetings for 4-H volunteers would fall into this category.

Agents often operating in the service area are working on building relationships and addressing basic needs of their clientele (Franz & Townson, 2008). Agents often start in this area when they are new to their position to build relationships and networks. Serving on local advisory councils or helping 4-H members enroll for the fair are examples of service. In this quadrant, an agent can often get evaluative results by the group meeting consensus or subsequent action taking place. There is often no formal evaluation used for this quadrant as the change is seen over time through systems changes of community indicators (Patton, 2008b).

Content transmission is a great way to meet the needs of a large section of the community. In this area, clientele often access the information directly, with little to no interaction with an agent (Franz & Townson, 2008). Examples of this include: fact sheets, social media posts, or a county website. Meeting outcomes and new practices being adopted show agents working in this quadrant the evaluative information they need (Patton, 2008b).

The final area, transformative education, focuses on changing social, environmental, and/or economic conditions (Franz & Townson, 2008). Through this process, learners “gain new competencies, apply what they’ve learned to their personal context, share results with each other, and adjust their application of learning as needed” (Franz & Townson, 2008, p. 11). Often these are the programs that are highlighted in reporting. Programs in this quadrant result in system changes and have a long-term impact on not just those that attended the program, but also on the communities in which that is needed. (Patton, 2008b).

NDSU Extension encourages agents to ensure that the majority of educational and programming time is spent in the high process, high content area of transformative education. While the other areas are important and have their place within Extension, the transformative education box is where the magic happens. Franz and Townson (2008) described transformative education as “when Extension agents develop long-term educational relationships with clientele to focus on changes in their learning and behaviors” (p. 11).

Evaluation is further limited by the way that much of the data is collected. Post-program surveys are very common in Extension. While this approach will provide the agent with data, other forms of data collection (focus groups, interviews, observations etc.) are much less frequent (Franz & Townson, 2008). Agents and community members alike may struggle to see the necessity of program evaluation. Agents may have a tough time balancing between spending time on evaluating programs and time implementing and delivering programs. In some states, this is addressed by having evaluation practitioner support, in others, there may be a need for additional support and/or training of current staff (Franz & Townson, 2008).

2.5.5. Evaluation in Other States

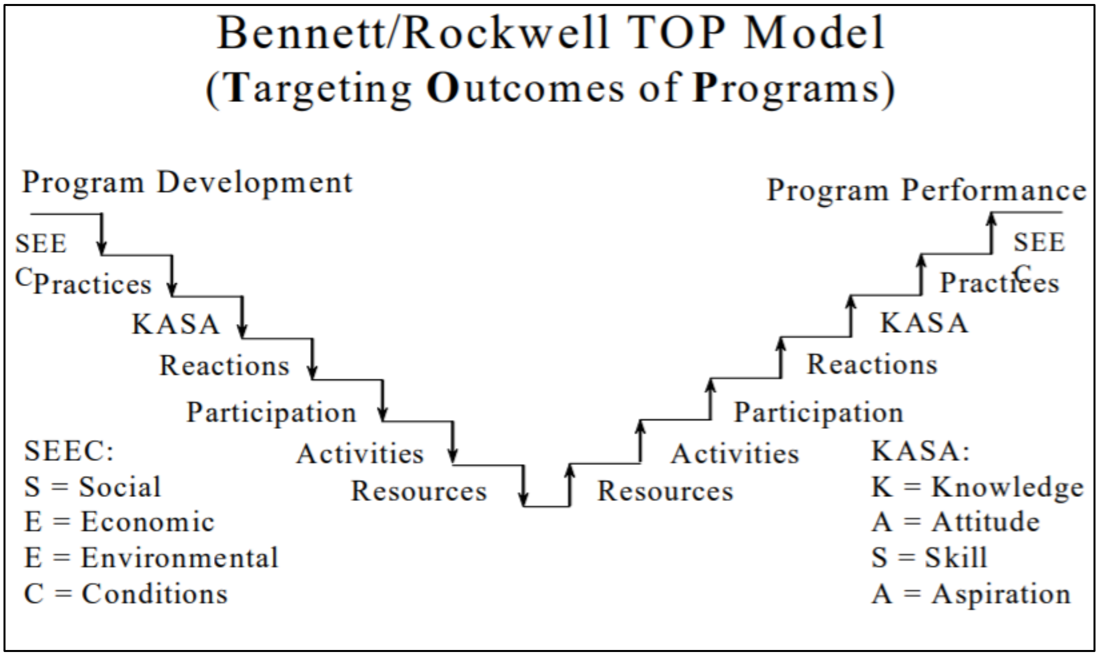
Evaluation across the Extension System varies widely. While some states have a heavy investment in the training and execution of evaluation, others rely heavily on the individual agents to design and implement the evaluation tool. Similar to North Dakota, Texas also uses Kirkpatrick's model to show impact (Texas A&M AgriLife Extension Service, 2019).

North Carolina Extension has a series of evaluation tools developed for agents to fill in information relative to their program. Evaluation tools vary based on the type of program or activity that is being evaluated. The tools are categorized into three major groups; one-time short training presentations lasting two hours or less, one-time relatively long training presentations, or multi-session training programs. North Carolina agents use the fill-in forms to evaluate programs they deliver (Jayaratne, 2013).

Tennessee Extension also has a specialist with a focus on evaluation. Impact statements focus on three levels of outcomes:

- Short term impacts that measure knowledge gained. This data is collected either immediately following the program or within three months.
- Intermediate impacts measure behavior change over a period of three months to one year.
- Long-term impacts take years to achieve and impact "social, economic or environmental conditions" (Donaldson, 2019, p. 6).

Arkansas utilizes the Bennett/Rockwell Targeting Outcomes of Programs (TOP) Model Hierarchy (Figure 2) for program, objective, and evaluation development. The TOP model differentiates between hard (ie. profit loss trends, direct observations) and soft (ie. self-reported change, staff recollections) data (Poling, 2005).



(Poling, 2005).
Figure 4. Bennett/Rockwell TOP (Targeting Outcomes of Programs) model.

This model walks agents through the process of creating a program and evaluation tool that will yield data that is then used to report the impact of a program or event (Mashburn, 2019).

2.5.6. Training on Evaluation

Ensuring that staff are properly trained on evaluation tools sets an organization up for success. Professionals at all points in their careers can benefit from professional development around evaluation. Stevahn, King, Ghere, and Minnema (2005) propose essential competencies for program evaluation as “a useful tool for identifying areas of existing strengths (e.g., particular evaluation approaches, methods, communication skills, interpersonal skills) as well as areas of need” (p. 46). These competencies would improve the field of evaluation through “(a) improved training, (b) enhanced reflective practice, (c) the advancement of research on evaluation, and (d) the potential for continued professionalization of the field” (p. 45). The six competencies are outlined in Table 4: Program Evaluation Competencies.

Table 3

Program Evaluation Competencies

Competency	Focus
Professional practice	Fundamental norms and values underlying evaluation practice, such as adhering to evaluation standards and ethics.
Systematic inquiry	Technical aspects of evaluation practice, such as design, data collection, analysis, interpretation, and reporting
Situational analysis	Analyzing and attending to the unique interests, issues, and contextual circumstances pertaining to any given evaluation
Project management	Nuts and bolts of conducting an evaluation, such as budgeting, coordinating resources, and supervising procedures
Reflective practice	One’s awareness of evaluation expertise and needs for growth, including knowing oneself as an evaluator, assessing personal needs for enhanced practice, and engaging in professional development toward that goal
Interpersonal competence	People skills used in conducting evaluation studies, such as communication, negotiation, conflict, collaboration, and cross-cultural skills

(Stevahn, King, Ghere, & Jane, 2005)

Patton (2008a) suggests one way to address several competencies is to set up a simulation in which the trainer uses potential results and walks the trainee through the whole process from start to finish. During this process it is important to ensure there is a shared definition and commitment to the process (Patton, 1997). Making the connection to evaluation and daily activities helps create a better understanding of evaluation and how it can be used to improve program reporting (Bailey & Deen, 2002).

Developing an evaluation tool can be a daunting task for agents. To avoid biases in the evaluation tool, agents need to look at questions that address:

- Overall feeling of the program (focus, format, relevance)

- Program dose (depth, frequency, amount/quality)
- Logistics and marketing
- Optional areas to cover
- Instructor and/or topic ratings
- Demographic information
- Needs for stakeholder reports (Roucan-Kane, 2008).

Ensuring these areas are addressed will help agents report efficiently and effectively. If a pre/post format will be used, it is recommended that the pre-program behavior be collected after the class as studies have shown participants are more likely to be truthful (Roucan-Kane, 2008). By administering the survey at the end of the training, participants have a better understanding of what the questions are asking and are less likely to inflate answers to what they feel is desired by the instructor (Raidl, et al., 2004). However, training agents is only part of the puzzle of increasing evaluation effectiveness and implementation. Ensuring that the conversation takes place with a broader group of organizational administration and decision makers ensures that the evaluation process and organizational learning are strong (Braverman & Arnold, 2008).

2.5.7. Special Considerations

As with the implementation of any statewide program, it is important to note any instances that require special considerations. With Common Measures two of these considerations are working in Indian Country and evaluating youth. To ensure anonymity, it is not disclosed if a participant in the study works in Indian Country. Notwithstanding, it is important to note special considerations that would need to be considered in certain areas of the state.

2.5.7.1. Evaluation in Indian Country

According to North Dakota Indian Affairs (2010), approximately 4.9% of the population of North Dakota identifies as Native American, 60% of whom live on one of North Dakota's five reservations (North Dakota Indian Affairs Commission, 2010). Therefore, it is important to understand how evaluations can be successful for this population. As Martinez, Running Wolf, BigFoot, Randal, and Villegas (2018) reported, conducting evaluations in Indian country can present its own set of challenges. "A contextual acknowledging of the impact of intergenerational trauma, supporting tribal sovereignty, and telling an indigenous story of practice development are critical elements of good evaluation practice in tribal communities" (Martinez, et al., 2018, p. 35). The need to involve those being evaluated is a critical step while developing evaluation tools for Indian County (Martinez, et al., 2018).

Further consideration needs to be given to how the data is collected. Martinez et al. (2018) strongly discourage the one-size-fits-all approach to evaluation, but rather encourage a community-driven process that takes the community into consideration. Having a "culture -and-community based evaluation" will contribute to a more effective evaluation of programs in Indian country (Martinez, et al., p. 44, 2018). The Native American Center for Excellence (NACE) reinforces the importance of relationships and cultural understanding prior to any evaluation being conducted. "Failure to engage the community in the partnership and integrate traditional practices that are congruent with tribal culture, language, and values can set up the evaluation for failure" (Native American Center for Excellence, 2019, p. 2). Making a clear connection between the data that is being collected and the community benefits shows the community that the researcher is there for more than just collecting data (Native American Center for Excellence, 2019).

Working with and evaluating members of the Native American population is a discussion that is needed for all agents in the state. Forty percent of the state's Native American population live in communities not on a reservation. It is critical that Extension agents understand cultural norms of the state's largest minority. The process of data collection that is utilized for Common Measures, a one size fits all survey, is a stark juxtaposition to what works well in Indian Country and should be considered with future statewide data collection initiatives.

2.5.7.2. Evaluation of Youth

By using artifacts, an agent can collect data that is nonreactive and is grounded in context of the program. Worker (2019) evaluated a youth program through qualitative means rather than the traditional quantitative means. Worker (2019) examined four criteria when evaluating his artifacts from a junk drawer robotics event:

- Complexity - How many different parts were used?
- Innovation - How did the youth modify the material to accomplish the end goal?
- Functionality - Did it work the way they wanted it to?
- Resolution - Did it solve the problem the group was given?

These criteria can all be linked to important life skills and competencies within the 4-H youth development field (resiliency, wise use of resources, planning, critical thinking, problem solving, and decision making). While this approach yielded a lot of rich data, and the approach was beneficial, it took a lot of effort to collect, analyze, and interpret the qualitative data (Worker, 2019).

Youth can be complicated to evaluate for agents. Often agents find themselves stretched too thin and as a result the time dedicated to evaluation is limited. By making program evaluation part of an agent's routine work, youth development organizations are able to get the information they need about a program without it being a burden. There is great power in including the youth as program experts throughout the process.

Due to the protected nature of participants in this study, it cannot be confirmed whether one of the participants is based in Indian Country. However, evaluating programs in Indian country and evaluating programs with youth have some commonalities that should not be overlooked by agents across the state. It is vital that agents examine evaluation practices to find possible opportunities in their counties for integration of these practices.

3. METHODS

3.1. Purpose of The Study

This research will examine the implementation of the Common Measures assessment tool by North Dakota Extension Agents through a qualitative review of the current use of the tool, the benefits and drawbacks of Common Measures, as well as ways North Dakota Extension and Common Measures can be improved moving forward. Areas for professional development related to evaluation and reporting will assist North Dakota Extension with future state side initiatives.

3.2. Research Questions

Five research questions framed this study:

1. Prior to Common Measures, how did Extension agents evaluate their 4-H program?
2. What are Extension agents' perceptions of the importance of evaluation in their role as an agent?
3. Do Extension agents view new programs and/or procedures, such as Common Measures, as intrinsic or extrinsic?
4. What training or professional development is needed in North Dakota to improve evaluation of the 4-H program?
5. How does an agent's length of service with Extension, position, educational level, and county composition influence differing viewpoints of evaluation?

3.3. Research Design

The study conducted using a basic qualitative research approach, intending to understand the meaning of an involved phenomenon. When conducting a basic qualitative study, the researcher is "interested in 1) how people interpret their experiences, 2) how they construct their worlds, and 3) what meaning they attribute to their experiences. The overall purpose is to understand how people make sense of their lives and their experiences" (Merriam, 2009, p. 23). A basic qualitative study focuses on analyzing data using recurring patterns that characterize data, and help make meaning of the interviews and data collected. While making meaning of data is at the core of all qualitative research, uncovering and interpreting these meanings is the "primary goal of basic qualitative research" (Merriam, 2009, p. 24). Basic qualitative research goes beyond uncovering a participant's experiences and attempts to understand the meaning behind those experiences (Worthington, 2019).

3.4. Population and Subject Selections

A purposeful sample was used to determine the participants in this study. "The logic and power of purposeful sampling lies in selecting information rich cases for in-depth study" (Patton, 2015, p. 264). By using a purposeful sample, there is a reduction in suspicion of why certain cases were selected. A purposeful sample allows the researcher to align "the inquiry's purpose, primary questions, and data being collected" (Patton, 2015, p. 264)

Administratively, the state of North Dakota is split into three Extension districts: west, central, and east (Figure 4) (NDSU Extension, 2019). Of the 53 counties, 20 are in the west, 20 are in the central, and 13 are in the east (Figure 4).

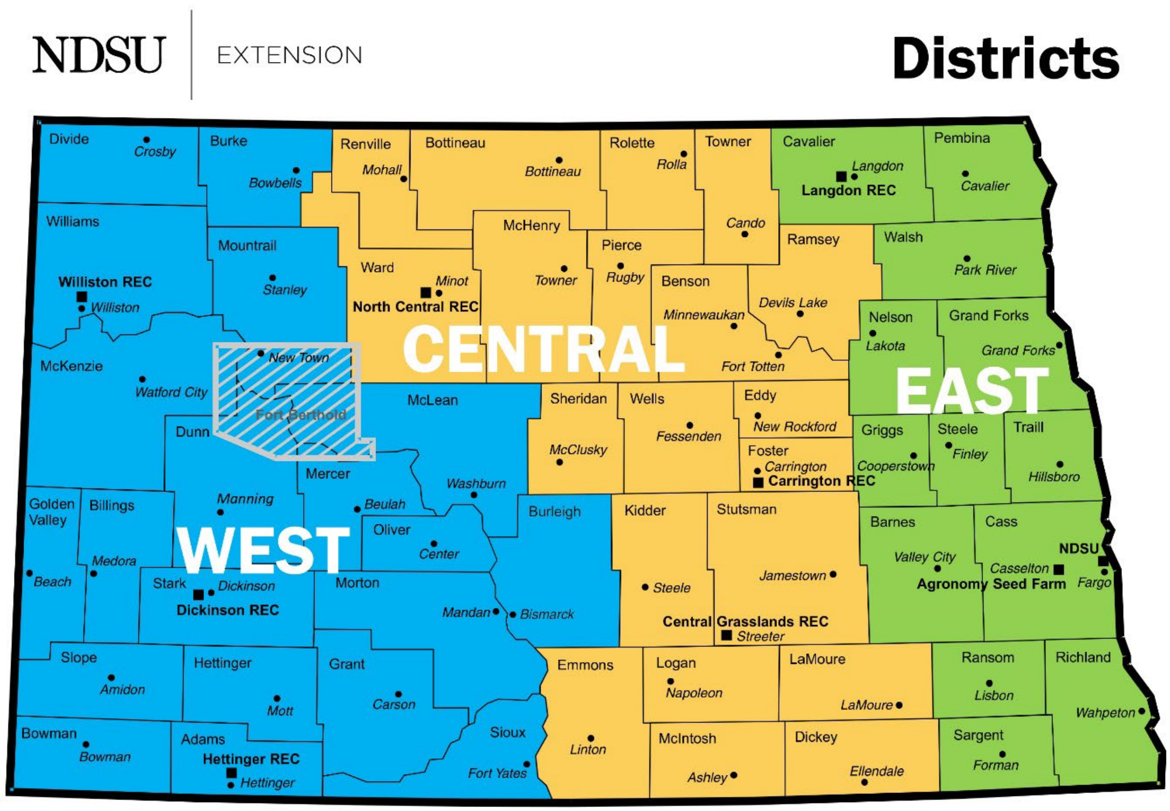


Figure 5. North Dakota Extension administrative districts.

While each Extension agent has a 4-H youth development responsibility, there is one agent in each county who serves as the lead for the 4-H program. In many counties, being the lead 4-H person is in addition to a position as a Family and Community Wellness (FCW) or Agriculture and Natural Resources (ANR) agent. Due to vacancies at the time the interviews were conducted, three counties did not have an agent in the lead 4-H role. For this reason, these counties were not considered for the study. North Dakota Extension agents (*N*=49) who held the primary responsibility for 4-H in their counties and were employed as November 9, 2018 were eligible to participate in this study.

Table 4: Youth participation in the 4-H program in North Dakota by county displays the total population as well as the youth population as reported by North Dakota Compass in

November 2018 (Wilder Research, 2018). 4-H membership numbers were collected from the statewide registration system and reflect 4-H enrollment of youth ages eight and older at the end of the 2016-2017 4-H year. These numbers were used to calculate the percentage of eligible youth in the county who participated in the 4-H program during the 2016-2017 4-H year. Counties were divided into three groups based on the percentage of youth participating, under 5%, between 5% and 10%, and over 10% youth participation. From each group, three counties were chosen, one representing each program area (4-H, FCW, and ANR) as well as each district represented. Urban or rural distinction as determined by the North Dakota department of commerce was also considered when determining participation. This yielded a sample size of nine counties. Due to one county declining to participate, an additional county was sent the recruitment email. The 4-H lead agent for these counties were sent the recruitment email (Appendix C) and were invited to participate in the study.

Table 4

Youth Participation in the 4-H program in North Dakota by County

County	Total Population	Youth Population	4-H membership (8+)	Percentage	Urban/ Rural	4-H Lead
1	181,516	38,869	339	0.87	Urban	4-H
2	95,273	21,792	134	0.61	Urban	4-H
3	70,770	14,522	125	0.86	Urban	4-H
4	67,744	16,382	211	1.29	Urban	4-H
5	35,350	9,388	175	1.86	Urban	FCW
6	31,095	7,144	301	4.21	Urban	4-H
7	31,916	7,995	182	2.26	Urban	ANR
8	20,917	4,281	137	3.20	Urban	ANR
9	16,239	3,524	67	1.90	Rural	FCW
10	14,301	4,982	51	1.02	Rural	FNP/FCW
11	13,632	3,840	144	3.75	Rural	FCW

Table 4. *Youth Participation in the 4-H program in North Dakota by County (continued)*

County	Total Population	Youth Population	4-H membership (8+)	Percentage	Urban/ Rural	4-H Lead
12	11,481	2,669	105	3.93	Urban	ANR
13	10,667	2,224	111	4.50	Rural	FCW
14	10,542	2,465	82	3.69	Urban	FCW
15	10,218	2,651	133	5.02	Rural	4-H
16	9,541	2,079	121	5.82	Rural	FCW
17	8,267	1,968	103	5.23	Rural	FCW
18	8,037	1,815	28	1.54	Rural	ANR
19	6,962	1,449	57	2.43	Rural	ANR
20	6,947	2,347	118	8.14	Rural	FCW
21	6,411	1,380	82	5.94	Rural	ANR
22	5,816	1,423	88	6.18	Rural	ANR
23	5,237	1,230	141	11.46	Rural	ANR
24	4,903	1,200	82	6.83	Rural	4-H
25	4,358	1,669	279	16.72	Rural	4-H
26	4,332	993	57	5.74	Rural	ANR
27	4,081	942	54	5.73	Urban	ANR
28	4,062	879	52	5.92	Rural	4-H
29	3,957	814	67	8.23	Rural	ANR
30	3,870	812	62	7.64	Rural	FCW
31	3,829	796	40	5.03	Rural	FCW
32	3,295	668	125	18.71	Rural	4-H/FCW
33	3,216	716	81	11.31	Rural	FCW
34	3,076	795	55	6.92	Rural	NONE
35	2,869	528	23	4.36	Rural	ANR
36	2,585	483	38	7.87	Rural	ANR
37	2,514	604	65	10.76	Rural	ANR
38	2,450	585	59	11.39	Rural	ANR
39	2,374	518	80	16.49	Rural	NONE
40	2,374	485	47	8.03	Rural	ANR
41	2,313	514	27	5.13	Rural	FCW
42	2,294	449	63	14.03	Rural	NONE
43	2,283	526	40	7.78	Rural	ANR
44	2,232	424	38	8.96	Rural	ANR

Table 4. *Youth Participation in the 4-H program in North Dakota by County (continued)*

County	Total Population	Youth Population	4-H membership (8+)	Percentage	Urban/ Rural	4-H Lead
46	2,100	560	28	500	Rural	ANR
47	1,952	434	64	14.07	Rural	ANR
48	1,903	401	31	7.14	Rural	ANR
49	1,903	431	61	15.21	Rural	ANR
50	1,769	455	41	9.51	Rural	ANR
51	1,349	241	21	8.71	Rural	ANR
52	763	164	59	38.98	Rural	ANR

(North Dakota Department of Commerce, 2017).

Of the ten invited participants, five (50%) completed the interview process. The agents recruited for the study formed a diverse sample based on county youth population, administrative district, position held, and population category.

Table 5

Research Participants

Participant Pseudonym	Title	Years of Service
Helena	4-H Youth Development	1.5
NPA	Family & Community Wellness	25
Lena	4-H Youth Development	5
Adeline	Agriculture & Natural Resources	4
Ellaine	4-H & Agriculture & Natural Resources	15

3.5. Instrument Development

Mills, Durepos, and Wiebe (2010) defined interviews as "conversations with purpose and direction" (p. 495). During a semi-structured interview, participants were asked to explain how they use any assessment or evaluation in their 4-H programs. Participants were also asked to describe how they view themselves as Extension agents and how they use and view assessment and evaluation in their programming. A semi-structured approach allows the researcher to both compare a participant's responses as well as to seek to fully understand each participant's unique experience (Mills etal., 2010). Table 7 lists the questions that were used as prompts to guide the interview. As is the case with guided interviews, there were instances that not all prompts were necessary to collect the needed information.

Table 6

Interview Questions

Questions	Probes
Your participation in this study will only be shared between myself and Dr. Adam Marx, my advisor. For this reason, in all reports you will be referred to using an alias. Do you have a preferred alias you would like to use?	
Tell me about you as an Extension Agent	What is your role within NDSU Extension?
	How long have you been with NDSU Extension?
	What kind of work did you do prior to coming to Extension
	How would you describe your 4-H program to someone outside of Extension?
	Do you see yourself as a good Extension Agent?
How do you define assessment?	What do you think your role as an Extension Agent is when it comes to assessment?
	How do you decide if a program or event was a success?
	How do you share this success with stakeholders and/or supervisors?
	Before Common Measures, how did you assess your programs?

Table 6. *Interview questions (continued).*

Questions	Probes
	If you could build your own assessment, in place of Common Measures from the ground up, what would that look like?
Now we are going to talk a little about Common Measures, when I ask questions about Common Measures, please think about any interactions you have had with Common Measures. Using only three words, how would you describe Common Measures?	Elaborate on each word chosen.
When you first found out about using Common Measures, how did you feel about the process?	Did your thoughts and feelings change over time? In what ways? What benefits do you see with using Common Measures? What drawbacks do you see with using Common Measures? Did you attend the Common Measures training in March? What impact, if any, do you think that had on how you see Common Measures?
Do you use Common Measures currently?	Why or why not? Why did you choose to do it this way? How have those in your 4-H community reacted to the new assessment? What do you plan to do with any of the data you have collected using Common Measures? - Do you have any reservations about doing this?
What do your supervisors or peers think of Common Measures?	Are you in line with these? Do you see Common Measures having any impact on the culture of assessment and evaluation in North Dakota 4-H?
What kind of community do you see among Extension Agents at NDSU?	How do you think these relationships affect your thoughts and feelings on programs and projects that are rolled out? How do you use this community to benefit your program?

Table 6. *Interview questions (continued).*

Questions	Probes
	Where do supervisors and administration fit in this community? Who do you see as the drivers of Extension? What positions do they hold?
Out of all the things we've talked about today in regard to Common Measures – or maybe some things we've missed – what should I pay most attention to? What should I think about when I read your interview?	

3.6. Data Collection

After creating the interview instruments and consent forms for participants, an application for permission to conduct the research was submitted to the Institutional Review Board (IRB) at North Dakota State University. The IRB approved documents are included in Appendices A through C.

Per the approved IRB, the ten potential participants were contacted via their institutional email to request their participation in the study. Included with the email were the interview questions as well as the informed consent form. The full text of the email invitation and reminder messages are provided in Appendix C. Interviews were conducted during November 2018 at a location chosen by the participants. They were asked to sign and send the consent form back to the researcher prior to the interview taking place.

3.7. Research Positionality

As this research is qualitative in nature, a discussion of reflexivity is critical. In order to adequately describe the researcher's beliefs, experiences, and expertise related to this study, it is necessary to shift to first person narrative. Since the credibility of any qualitative research study largely relies on the researcher, it is important that the reader understand my views, experiences, and knowledge in the areas of Extension and evaluation (Patton, 2015).

As of the time I am writing this, I have recently left my position as an Extension agent in Burleigh county North Dakota, a role I served in for four and a half years. I began my master's program six months after starting my position. During my time as an agent, I had three different supervisors, all with differing viewpoints on evaluation and assessment of an Extension program. After one and a half years into my time as an Extension agent, I was given the responsibility of being the county coordinator. In this role, I was responsible for administrative duties within the county office, including, but not limited to, supervising support staff, managing relationships with the county commission, and serving as the lead for larger office wide programs.

In an effort to show the value and impact of the 4-H program in Burleigh county and to meet the requirement of our annual review, I used much of what I learned in my graduate courses to try and develop an instrument that could be used to show this impact. Throughout this process, there continued to be issues with meeting the requirements of the Kirkpatrick model, and thus not meeting the requirements for my annual reports. My interactions with other agents across the state showed me that many had exceptional programs and were doing wonderful programs, but often did not choose to report on these programs as the evaluation tools were not available. While this is not an issue for all agents, as most agents in North Dakota have a dual appointment, they often have programs from their main field (FCW or ANR) that they can use to show impact for annual reporting requirements. Many of the packaged programs come with developed evaluation tools that require little to no adjustment by agents prior to them delivering the program.

Prior to my employment with Extension, I worked in the disability services field. In this field, evaluation is vital to ensuring people receive the services they need. Both formal and informal assessments are completed on a regular basis. There are many similarities between Extension and disability services, including the mandated nature of the assessments. However, one stark difference stuck out to me when I transitioned to Extension. It seemed that while evaluations were being conducted, they were not being used to their full potential to aid in program improvement and development.

In Extension, evaluation is pushed as something that is vital to show that our programs are of value to the people of North Dakota. Usually these evaluations are retrospective in nature and look something like this:

- Prior to attending this class, how comfortable did you feel confronting people with different personalities than you? (scale 1-5)
- After attending this class, how comfortable do you feel confronting people with different personalities than you? (scale 1-5)

From these questions, agents are able to gather solid number driven data. However, due to the nature of the 4-H program, these assessments do not always work the best for assessing 4-H programs. I often got creative with assessing programs and using observations to decide if programs were successful. Within the 4-H program, the biggest changes and impacts are often not seen right away, or even within the same year. Trying to show this impact makes the reporting process even more challenging. The issue with this is that an assessment looks very different for youth than it does for adults. I can reach out to adults six months after a program and collect additional data to show knowledge and behavior change. However, completing the same process for youth brings a new set of challenges, and is often not successful. Often this additional data is collected via an email or phone call, two points of contact that are often not as accessible with youth.

In September 2017, I began working with Dr. Meagan Scott and a committee of agents across the state to determine the parameters we would set for collection of the necessary data across the state. Shortly after we began the process of implementing the Common Measures study, I was approached by Dr. Scott to conduct research on how agents were using Common Measures in their county programs. As an agent, I had been told many times by a supervisor that the type of evaluations I was conducting on my programs was not meeting the requirements for impact statements. When the Common Measures tool was introduced, I had feelings of doubt, as I was not sure this tool could meet the requirements of my yearly reporting.

Aside from failing to see how the Common Measures assessment tool would meet the impact statement reporting requirements, I felt that the it was going to provide the state with some useful information that could show the impact 4-H was having across the state. However, I soon learned that many of my colleagues felt the same way. Many of them said they would administer the Common Measure instrument because it was required but would not use the data in any of their own reporting. The extrinsic drive to complete an evaluation seemed to be the main motivation for doing any evaluation.

In my position as a 4-H youth development agent I also had a partial state appointment, which was primarily working with the camping and civic engagement programs. We tried to use the Common Measures assessment to show impact of these programs. However, it was met with varying degrees of success. One of the biggest issues with Common Measures is the inability to connect it to an individual program that was completed. For example, I was part of a state wide citizenship and community service program that focused on leadership development and civic engagement. Common Measures was used to evaluate the program. Without data showing that the weekend was the reason the youth answered a certain way, it does not show that 4-H is the reason for the behavior, thoughts, or feelings, and thus cannot be used to show impact.

In my county program, the first time we were asked to use Common Measures we were asked to administer the survey to a percentage of youth in our county. To guide the roll out of Common Measures, a team of agents from across the state worked with Dr. Scott to develop a plan and reasonable expectations for all counties. During this collection round, I heard concerns from very few agents. I was able to administer the survey during a program that was already happening in my county, an approach of many agents. I saw the implementation process of this as a team effort and appreciated that those who would be collecting surveys were also involved in the planning.

In March 2018, another round of Common Measures was released to Extension agents. This time it was expected that surveys would be completed by August 2018. To me, this implementation was not well timed, as many agents were done, or close to done, with their 4-H programing for the year, as members started to prepare for the fair season. This time there was not an expectation that each county was responsible for a certain number of surveys, but rather, as a state, there needed to be 120 surveys collected in each of the areas of: citizenship, college and career, healthy living, and science.

3.8. Data Analysis

Data collected in the study was transcribed using Rev.com, a third-party transcription service. After transcription, a series of coding was performed on the data. Coding provides a way to "translate" data that is collected into meaningful codes that are then used for data analysis (Vogt, Vogt, Gardner, & Haeffeke, 2014, p. 13). A code "capture[s] a datum's primary content and essence" (Saldaña, 2016, p. 4). Coding takes the data collected through what Saldaña (2016) describes as a cyclical process. First cycle coding was completed using in vivo, initial, and versus coding. After each cycle of coding an analytical memo was written.

3.8.1. Initial Coding

During initial coding, "data are broken down into discrete parts, closely examined, and compared for similarities and differences" (Strauss & Corbin, 1998, p. 102)**.** Initial coding is the first step on a winding journey of further exploration. During initial coding it is imperative that the data be approached with an open mind, free of any bias. All codes from this cycle are "tentative and provisional" and as analysis continues, may be reworded (Saldaña, 2016).

Each interview was read thoroughly with a specific research question in mind and codes were assigned accordingly. After reading through all five interviews for each question an analytical memo was written. These memos served as a place for the researcher to reflect on emerging themes and nuances within the data that may be over looked when analyzing the data on a larger scale. After the first cycle of coding a few key themes emerged in the data:

- The length is an issue for many people.
- The vagueness of CM is leading people to not think that they can use it past collecting data for the state. Helena referred to the CM measure data as data collected for "them", them being the state. She called it "their data."
- A few agents (Helena, and Lena to an extent) see evaluation as time not well spent. They would rather use that time to be doing programming.
- LPA stated she does her impact statements on programs that already have an evaluation tool.

- It is important to be mindful of Adeline's feelings. She has some great things to say about the data and why it matters to collect. She didn't really have much negative to say about the tool and was thankful to have it provided.

3.8.2. In Vivo Coding

Similar to initial coding, in vivo coding is a first cycle coding that allows the researcher to connect with the data. However, one stark difference between the two lies within the codes that are used. In vivo coding focuses on the words used by the research participants, as this allows the research to "honor the participant's voice" (Saldaña, 2016, p. 106). By using the participants' voice, the researcher honors the meanings in the participants' experiences (Saldaña, 2016).

3.8.3. Versus Coding

To better identify the dichotomous relationship that can exist between an agent and outside demands (supervisor, funders, or partners), versus coding was used. "Discerning the conflicting power issues and micro-politics among constituents and stakeholders is an important diagnostic for initiating and facilitating positive social change" (Saldaña, 2016, p. 137). Versus coding allowed the researcher to identify intrinsic and extrinsic drives for completing evaluation as seen by participants. Versus coding leads to what Saldaña (2016) defines as "major moieties" (p.138). Versus coding also identifies the contributions to the conflict and how, if at all, the conflict can be resolved. Versus coding does not aim to find a "clear-cut hero and villain," but rather brings to light the power issues at hand, which are often seen by humans as dichotomies (Saldaña, 2016, p. 140).

In this study the following versus codes emerged: good programming vs. good evaluation, agent vs. agent, state vs. county, transformational change vs. Common Measures, verbal survey vs. paper survey, local needs vs. reporting needs, us vs. them, and culture vs. requirements.

3.9. Second Cycle Coding

Second cycle coding focuses on developing a sense of “categorical, thematic, conceptual, and/or theoretical organization from your array of first cycle codes” (Saldaña, 2016, p. 234). Second cycle coding takes the codes from the first cycle and synthesizes them into more succinct themes and concepts. For the study, a second cycle coding of pattern coding was completed.

3.9.1. Pattern Coding

Pattern coding helps to make sense of the first cycle codes. Saldaña (2016) defines pattern codes as “explanatory or inferential codes, ones that identify an emergent theme, configuration, or explanation” (p.236). In a sense, pattern coding is the development of super codes that help further synthesize the data into themes that help to answer the research questions (Saldaña, 2016).

During pattern coding, each participant was assigned a specific color of sticky note. Every code that was developed during first cycle coding was written on a sticky note and was placed on a poster board reflecting the research question the code helped answer. Once all of the codes were distributed to the research questions they were grouped into themes. Some of the themes from this cycle of coding were: meeting requirements, evaluation being an afterthought, qualitative data collection, program planning, working with diverse audiences, and changing perceptions of the 4-H program.

3.10. Ensuring Fairness

A debate remains if the aim of qualitative research should be objectivity or trustworthiness (Patton, 2015). While both sides have compelling arguments, neither can be achieved perfectly. For the study, fairness will be used as indicators of non-biased results.

Fairness has four defining features:

- Acknowledgment and equal representation of both sides of the case
- Equal representation of each side of the case to ensure the reader has all the facts to draw independent conclusions
- Ensuring that biases are checked during an interview and have limited involvement in the presentation of facts
- "A relative criterion that is measured by balance rather than by isomorphism to enduring truth" (Patton, 2015, p. 725)

4. FINDINGS

Of the five interviews, three were conducted in person and two via Skype. Overall, agents were very open to discussing the pros and cons of Common Measures and evaluation in North Dakota. After each interview a memo was written discussing the interview and any themes or points of interest were pointed out. These memos were later used during coding to ensure key findings were not over looked.

The findings from the study are presented as they related to each of the proposed research questions. It is important to note that several of the themes intermixed between the questions and evidence of multiple questions can be found in each of the following sections. The use of pattern coding allowed the researcher to synthesize the codes into the main ideas of the research questions. In addition to themes emerging to answer the proposed research questions, additional information regarding the culture of North Dakota Extension and the interplay between counties and agents emerged. These findings are presented in section 4.6.

4.1. Prior to Common Measures, How Did Extension Agents Evaluate Their 4-H program?

Three major themes formed the data for this question, provided surveys, qualitative data collection, and questionnaire. Responses from those interviews told a story largely of evaluation being required and often difficult for an agent to execute fully independently. While agents are collecting qualitative artifacts, they lack the basic skills to use that data to tell the story of their 4-H program.

Overwhelmingly, agents interviewed for the study used qualitative data they collected through a variety of means to determine the success of their 4-H program. Surprisingly, none of this data was used to write any impact statements. All the impact statements gathered and discussed with the participants were focused on a single educational program, rather than the whole county 4-H program. Agents stated they knew their programs were a success after examining items such as, a waiting list, public response to the program, and publicity in the local media. This type of data collection allows agents to determine if their program was successful and what needs to be done to improve it in the future. However, the fault in this form of data collection is the disconnect from the state mandated reporting requirements. “I can tell a story, but I can’t write a great impact statement,” explained Ellaine, “I can tell you what we’re doing, but telling it and putting it on paper, you lose a lot of the passion.” When asked why they had not used this data to tell the story of their program in an impact statement, four of the respondents stated they did not feel it met the level of evaluation that was required from the state (3-4 on the Kirkpatrick model). The other agent stated there were different programs that rose to the top every year that she felt were more worthy of reporting.

Two of the agents, who both have a split appointment, stated that the impact statements they have chosen to write over the years are related to surveys that have been developed by specialists and come as part of a packaged program. One participant stated that the reason that she has never written one on a 4-H program due to the lack of packaged programs with valuable surveys included. NPA stated it had been a while since she did an impact statement on just 4-H. When asked why, NPA stated that other projects and programs that provided survey tools “were far better at gathering information. I hate having to develop a survey by myself. And it’s just easier because some of the other programs that we do, there’s already [an] evaluation tool that’s provided for us.” NPA later added that building an evaluation makes her “lose time.” Finding this balance between a quality program and a quality evaluation tool was expressed by several participants. The struggle to meet the needs of state reports as well as continue to meet the needs of the county were very apparent.

While agents are evaluating and collecting data on their 4-H program, they are not taking the next step to report these findings and help tell the story of the 4-H program. A divide between the expectations of reporting and the tools agents are given to gather and report the data continues to grow as reporting requirements for the state and program partners evolve. This divide has caused agents to not tell the whole 4-H story.

4.2. What Are Extension Agents' Perceptions of the Importance of Evaluation in Their Role as an Agent?

As Bailey and Deen (2002) discuss, a basic understanding of the tools being used and how they can directly impact an agent's programming is critical. Not only does an agent need to have a good understanding of how to use the tool, it is also important that a clear understanding of how the information collected can help benefit their position and the work they are doing every day.

"An assessment of learning, I guess I see that more as a test or actually seeing visual results. Whereas an evaluation is, I guess in our world, you intend to make change," said Adeline. Showing an intent to make change on the Kirkpatrick scale is at the center of how agents believe they need to report impacts. While most agents understand the importance of evaluation as an agent, not all see it as an integral part of the position. Helena described evaluation as bean counting, "I feel like there's so much bean counting and if we spend all of these hours recording all of the people we see, it's taking away from time that we could be using for our job." The formality of reports and impact statements is where agents tend to start seeing the evaluation as more of a requirement and less of a benefit to their program.

The most common way data are collected by agents for reporting purposes is a survey after the program. In some instances, these surveys are done before participants leave the program, and in others, are done as a follow up several months after completion of the program. Data collected this way can be used to develop statistical reports in support of the program. Yet, often data are not valuable to an agent to evolve the program. “I feel like a part of us needs to appease the state, and the other part of us needs to appease our county. I would rather know how the program is doing,” said Helena who further elaborated,

> what I can do better for the people in my county. And what programs they’re interested in and want to learn more about. I guess, I am more concerned with what they think and feel, but I think sometimes the state has certain things they want us to find out, and then I don’t want to say push their agenda, but.

Agents often see the evaluation tool to meet the reporting guidelines as an afterthought, or something that is pushed aside due to lack of time. All the participants in the study stated that they gather data to determine the success of a program through informal qualitative means. These data are used to help refine programs for the future and adjust as needed but the information is not used to report on the successes of the program through formal means of an impact statement. While the agents all mentioned collecting this data, they did not refer to it as data or evaluation but rather stated it was how they personally gauge success of their program.

When asked for her thoughts on Common Measures and the importance of assessment Ellaine said;

> I think it [assessment] is important, but I think it has to be done so it’s effective for the communities we serve. When I started there was a CYFAR grant, and the pre and post-test were both 35 pages long. It was rather disastrous because I don’t think the information they received back was accurate because people were burned out. A lot of

> the questions weren't relevant to the population. They got data, but I don't think it was accurate. I think if you give data, it needs to be accurate.

Similar to Ellaine, other participants see the importance of the data being specific to the county they are serving to help tell the story. Collecting data through the administration of a survey does not work for all populations that agents in North Dakota serve. "Literacy skills are not always the best by some of our older members and parents of our youth…so we do verbal, and I gather data that way" stated Ellaine. It is important to note, that Ellaine gathers all her data this way. If a survey is requested for a specific program, she will use oral assessment to collect the data that is necessary to meet the reporting requirements.

Extension helps shape the narrative of North Dakota, especially in the rural areas of the state. Three of the participants stated that they feel evaluations provide them with the information they need to better tell the story of the 4-H program. As Adeline, who is from a rural county stated, rural counties are often not represented in the big picture of the program. Data from surveys, such as Common Measures, gives her the resources to be able to say, "hey, we are here. And we have some amazing youth in the county, and they have the potential to do whatever they set their mind to." The narrative of the importance of rural youth and families is not lost on agents, and 4-H allows Extension to have an increased connection with these important stakeholders. Leana described 4-H as the "gateway to youth."

4.3. Do Extension Agents View New Programs and/or Procedures, Such as Common Measures, as Intrinsic or Extrinsic?

Overwhelmingly, agents felt that Common Measures was extrinsically motivated. "I didn't want to mess up their data" said Lena when discussing how she plans to use the Common Measures data to tell the story of her county program.

> I have pulled from blocks of questions in the past, when I've tried to evaluate things, but I kind of feel like what you just said, it doesn't get at the beef of what the stakeholders might want to know.

From a broader perspective, agents felt that reporting and evaluation was mandatory and done to appease the state. Interestingly, after further conversations, agents reported that they do evaluation on a regular basis of their program, however, it is informal in nature and does not yield the necessary data that is required for reporting. However, every agent interviewed said they use alternative forms of data collection (school district collected data, enrollment numbers, focus groups) to gather data about the health of their program and what could be done better. This data are often not viewed as data that are good enough to report findings to the state. When asked about evaluation, none of the agents mentioned this type of data collection, even though all of them are collecting data in these ways. While participants see the value in using existing data to tell the story of their 4-H program, they lack the understanding and skill to transfer that to the necessary reports for the state. All of those interviewed stated that qualitative information helps them better address the needs of their county.

Much of the data that agents are informally collecting helps paint the picture and tell the story of their county 4-H program. Ellaine shared a success story of a youth in her county who was the first in her family to go to school beyond high school who recently graduated with her master's degree. As Ellaine put it, "those are the kind of impact statements I think we need to be getting out there." The data collected for this impact were not gathered with a survey. The impact that was not achieved in a year, or even five, it took years of programming and development of the young person to see the results. These are the stories that agents tell, however these same stories are missing from the state required impact statements.

A few agents reported that their clubs and members understood the importance of completing evaluations, and, while they didn't like it, were willing to complete the surveys if it helped the 4-H program. NPA described Common Measures as "cumbersome" and "missing the mark" on several items. "That's Common Measures, it's a long survey, it's a long evaluation tool. Do these young people, do they have the patience with their attention space to sit through it and finish it? Are they just checking boxes?" stated NPA.

On the other side of the spectrum, when Helena was asked if she had a better understanding of the questions and the data they provided if she would use Common Measures she expressed concerns over having a survey for everything. As a follow-up Helena was asked if she felt Common Measures could replace some of the surveys she is currently using, she said

> I wasn't super impressed with the questions that they had listed or if it would pertain to the program that I would do […] I think it's great and it'd be a good indication […] that kids do like 4-H but I don't think it explains why they like 4-H so much.

Overwhelmingly, agents lack the knowledge and skills necessary to take the information they collect without a formal survey and put together a report that tells the story of their 4-H program. Because of the lack of connection between non survey collected data and reports, agents see evaluation as checking a box, completing a requirement from the state, and extrinsically motivated.

4.4. What Training or Professional Development is Needed in North Dakota to Improve Evaluation of the 4-H Program?

To create a better connection between the data agents are collecting to improve their program and the state mandated reports, agents across the state would benefit from additional training and resources. Training in the areas of qualitative data collection, analysis and reporting, background of questions asked on Common Measures, developing an evaluation tool from scratch, and support after trainings were common themes throughout the interviews.

Agents are collecting qualitative data to personally gauge the success of the program. However, they are not using that data to report on the impact of their program. The research indicated that agents lack the knowledge necessary to use data collected this way to report program successes. While Common Measures is quantitative in nature, agents feel that a story is needed to accompany the numbers. Ellaine explained,

> if one of my kids [4-H members] doesn't understand a science project that we're doing, I want to know why they're not understanding it. I just don't want to say, 'do you like science?' because to me, there's no depth to that. I like things in depth. If I'm going to do an evaluation, I want to get some real hard-core data.

Agents have the quantitative data and stories. However, they are not equipped with the training and tools necessary to make the connection to the reporting required by the state.

On the issue of meeting the reporting requirements, agents felt that more training was needed on how to develop a survey tool that reaches this goal as well. Lena stated it would be very beneficial if a specialist "took a program and then used Common Measures data to write an impact statement." This approach would show the agent "this is how you can use the data effectively." Several participants stressed the disconnect between Common Measures and the state required impact statements. Agents need better direction on how to connect the two.

Additionally, district directors and Extension leadership would benefit from training on Common Measures and the 4-H program. Ensuring that leadership and agents are functioning with the same knowledge will create a better understanding of 4-H program evaluation on all levels of the organization. While no agent interviewed stated their district director was against

Common Measures, Helena said her district director wanted to hear stories, “I mean, I had my impact statement, and he barely looked at it, and just wanted to know stories, and then how many kids were enrolled.”

Participants did not feel as if they had enough knowledge on the background of the questions on Common Measures to fully understand the entire project. Even though Helena attended the state-wide training on Common Measures, she commented that she was “not super well-versed in it [Common Measures].” Specifically, questions such as, “how many hours do you typically spend on 4-H activities each week” (National 4-H Council, 2018, p. 9) which are explained by research, do not resonate with agents. By providing training and background information more specific to the tool, may help agents understand the rationale behind the questions asked.

Both agents that attended the training sponsored by the state stated that upon returning to their counties they felt like they did not have the knowledge needed to successfully implement the survey successfully in their counties. Better tools and support once agents have left a training and are back in their counties is imperative to program success. Additionally, providing tools and resources to help agents get through different situations that may arise is necessary. As Ellaine stated, “make a program fit the people you serve.” Knowing how, and when you can manipulate the tool to get better data that not only gets the data that is needed need for the state and national reports, but also provides the local program with information that will contribute to the progress and success of the program. Specifically, it was expressed that working with diverse populations such as kids not enrolled in a 4-H club (but participating in Extension youth development activities), and younger youth (Cloverbud age) were not addressed by the tool and create barriers that agents must overcome to be successful in collecting data.

Participants expressed concerns with needing to find a balance between delivering high quality content and having a high-quality evaluation. Through a wider lens, participants expressed the need for more planning and collaboration sessions that help achieve the goal of the evaluation being a part of the planning process rather than an afterthought. States that have a dedicated evaluation specialist were brought up as a way that would provide beneficial support to agents across the state. Adeline explained,

> Don't just say, 'Well, here it is.' Teach me how do I know what to ask? Or what's the data we're looking for? I can think in my head, like, 'Okay, I know I want this.' But then, jeepers. I get the results back and I'm like, 'What the heck? This is nothing. I should have asked this.' Just being able to help think through that process.

Agents need help not only forming the evaluation, but also analyzing the data so it is valuable for required state reports as well as program development and improvement.

4.5. How Does an Agent's Length of Service with Extension, Position, Education Level, and County Composition Influence Differing Viewpoints of Evaluation?

The agents in the study came from a wide variety of previous experiences, education, and lengths of service to Extension. By better understanding what helps an agent feel more connected with evaluation, Extension can provide professional development opportunities that are more targeted to an agent's specific needs.

Overall, an agent's length of service had a positive influence on their views of evaluation. Helena, who has been with Extension for the shortest period had very negative thoughts and feelings about any formal evaluation tool. Formal education in Extension did not seem to coincide with any of the participant's thoughts on Common Measures. It was often participants with no formal educational background in education or Extension that saw the most value in a standardized evaluation tool.

NPA, whom has 25 years of experience stressed the importance of an evaluation tool, like Common Measures, on the 4-H program across the state. She was supportive of the idea of the tool, however, did not feel that Common Measures answers the questions that needed to be answered to achieve the result that was trying to be achieved. “I was glad that we finally had a formal tool, when I looked at the tool, you know, some of the questions just didn’t quite seem to hit the mark.”

Ellaine has been an agent in two counties and with NDSU Extension for fifteen years. Ellaine described herself as

> a pretty good follower. If they say ... Because I'm a believer even with grants, when you sign up for a grant and it tells you, "These are the expectations," you signed up for it. You need to do it. That's the same way with Common Measures. If that's what's expected, I'll do it. I'm a good follower.

Ellaine stated she would do the survey because it was asked of her but would more than likely not use the data for anything she needed.

Lena, who has five years of experience in Extension stated, “I think it [evaluation] is a really important part of being an Extension Agent that I didn’t truly appreciate or fully understand until more recently.” Lena stated throughout her interview that much of the data she uses to define a successful program comes from conversations with people in and around her county.

Adeline, started her current position right after college and has four years of experience in Extension. At the time of the interview, agents had not received their county specific data back, however, Adeline was hopeful for the results,

> I would like to show our commissioners, really, the impact 4H has had. Because like I said, we had 25 kids. When that Common Measures survey came out, that's when we were really growing. So it'll be really neat to see. I just really want to share with commission and general public. That's not state stats. That's [county name removed] County statistics, and that's so important.

Helena, who has been in her position a little over 18 months at the time of the interviews was a bit more skeptical about Common Measures, "I guess the brief amount of Common Measures that I saw, I wasn't super impressed with the questions that they had listed or if it would pertain to the program that I would do."

All participants in this study are considered the 4-H lead for the county they work in. However, those that hold a split role often did not focus their impact statements on the 4-H program. NPA and Lena both stated that the ease of using a premade evaluation tool that is available with packaged programs helps them determine the program they will formally report on. Due to the nature of the 4-H program, especially in North Dakota, packaged programs are very limited, and thus, 4-H agents are not able to use a pre-made evaluation tool and must develop one of their own.

Due to the small sample size and limited areas considered urban in North Dakota, no conclusions could be drawn on the difference between rural and urban counties as it related to evaluation.

4.6. Additional Themes

Throughout the research, themes not related to the research questions emerged in the data that are worth noting. Every participant in the study noted a county vs. county competition that exists in North Dakota. Several times, this was brought up in a negative way and as something that needs to be addressed in the state for Extension to continue to move forward. Adeline referenced a former document that was often updated and distributed within Extension that color coded counties based on membership changes over the last 5-10 years.

> It was tough when [the state office] would make that freaking map of North Dakota. To look at my county being red… Red. Red, red, red. Why is it the amount of kids that matters? Why isn't it the experiences that we're giving them? That can be different, it is different within every county.

The red Adeline is referring to noted counties that had not seen a membership increase over a certain amount. Adeline continued by saying that she is hopeful Common Measures will help change the narrative from county vs. county to more experiences based. While a bit of competition is healthy, it should be noted that several participants stated they feel the competition often hinders program development across county lines as well as program areas.

As Lena said, "I think there is a disconnect about what 4-H is, even internally." Agents with a split role described this as more of a balancing act, a constant push and pull between 4-H and their other program area. Even though every agent is required to have a role in 4-H youth development, counties with multiple agents thinks this looks different if they are not the 4-H lead for the county. Lena said that NDSU Extension could benefit from more cross-programmatic programming,

> there already has been a lot of focus on cross-programmatic planning and I think the more from the top we are encouraged to use our specialists in other disciplines, partner, I really think the more we partner, even as counties, on things, we build these relationships that open us up to new resources. And I kind of think a lot of that is some counties already do it so well, so strongly. They help each other out all the time and other counties have their distinct disciplines and so I guess the more it's encouraged, that community will become stronger.

Cross-programmatic planning would hopefully alleviate some of the county competition that agents reported during interviews.

4.7. Findings Conclusions

Overall, the agents that participated in the study were initially hopeful for Common Measures and the data it could provide stakeholders at all levels of the organization. However, after the initial implementation of the survey, agents feel it could use some additional refinement. In addition to refinement it is important to note the needs that arose out of the research for additional training and resources. Given the right tools, agents are willing and able to use evaluation and assessment to address reporting requirements as well as local needs.

5. JOURNAL ARTICLE FOR THE JOURNAL OF HUMAN SCIENCE AND EXTENSION

Per discussions with my advisor and committee members, a journal article was written in place of chapters five and six.

5.1. Abstract

Using a basic qualitative approach, interviews were conducted with five agents in North Dakota with a range of experiences and lengths of service to evaluate the effectiveness and implementation of Common Measures. The study addresses the culture of North Dakota Extension's views on evaluation and the implementation of Common Measures. The use of a statewide reporting tool to assess the 4-H program in North Dakota was perceived as necessary, however, many agents felt the tool, Common Measures, missed the mark and did not meet the requirements that exist in North Dakota Extension for report. The ability to tell the story of the 4-H program and the long-term impact it has on youth is necessary to the success of any statewide reporting venture. More work is needed to help train agents on how qualitative forms of data collection can be used to tell the story of their program.

5.1.1. Key words

Common Measures, evaluation, 4-H, Extension

5.2. Introduction

The Extension service has a history dating back to the mid-1800s, of addressing community needs through hands-on trainings and programs. While the format of reporting impacts of such trainings and programs has changed, the importance of reporting is still vital to the success of Extension. In 2002 Lerner and Lerner (2013) with Tufts University started a longitudinal study that followed youth participating in the 4-H program, as well as those who were not, for 10 years to see the impact that the organization had on their development. The

Lerner study contributed greatly to the development and advancement of the positive youth development field. However, with the research approaching 10 years, it was time for the field to find new data and research to show the impacts of the 4-H program. This process started in 2012 with the development of Common Measures 1.0 with version 2.0 following in 2018 (National 4-H Council, 2015). Common Measures, as it stands, “is designed for youth in fourth through twelfth grades who have participated in a 4-H program project, or grant” for a minimum of six contact hours (Hawley, 2017, pp. 2-3).

Within the North Dakota 4-H program, Extension agents from across the state were tasked with administering the survey to a minimum of 30% of the 4-H youth in their county. The enrollment in North Dakota 4-H as of August 2018 was 4,810 youth, making the target of 30% of youth surveyed, 1,443. At the completion of the project 2034 surveys were collected across the state. The initial wave of surveys used the universal tool. Youth received the survey at a variety of places including, county fairs, awards nights, and club meetings. In addition to county-based programming, state specialists also used it for statewide programming.

Currently, NDSU Extension uses the Kirkpatrick model to determine impacts of programing within Extension. The state requires each agent to evaluate at least one program annually (NDSU, 2008a). The Kirkpatrick model splits evaluation into four levels of impact; reaction, learning, behavior, and results (Kirkpatrick Partners, 2018). Using this model, agents develop tools that are used to show the impact of their programs. As evaluations move from level one to four, the impacts of a program become more impactful and meaningful as well as more expensive, complicated, and time consuming (Kirkpatrick, 1996). As a requirement for reporting, agents are asked to submit an impact statement annually that shows a level 2.5 impact or higher (NDSU Extension, 2008).

5.2.1. Purpose of The Study

This research will examine the implementation of the Common Measures assessment tool by North Dakota Extension Agents through a qualitative review of the current use of the tool, the benefits and drawbacks of Common Measures, as well as ways North Dakota Extension and Common Measures can be improved moving forward. Areas for professional development related to evaluation and reporting will assist North Dakota Extension with future state side initiatives.

5.3. Methods

Five research questions framed this study:

1. Prior to Common Measures, how did Extension agents evaluate their 4-H program?
2. What are Extension agents' perceptions of the importance of evaluation in their role as an agent?
3. Do Extension agents view new programs and/or procedures, such as Common Measures, as intrinsic or extrinsic?
4. What training or professional development is needed in North Dakota to improve evaluation of the 4-H program?
5. How does an agent's length of service with Extension, position, education level, and county composition influence differing viewpoints of evaluation?

5.3.1. Research Design

The study was done using a basic qualitative research approach, intended to understand the meaning of an involved phenomenon. When conducting a basic qualitative study, the researcher is "interested in; 1) how people interpret their experiences, 2) how they construct their worlds, and 3) what meaning they attribute to their experiences. The overall purpose is to understand how people make sense of their lives and their experiences" (Merriam, 2009, p. 23). A basic qualitative study focuses on analyzing data using recurring patterns that characterize data, and help make meaning of the interviews and data collected. While making meaning of data is at the core of all qualitative research, uncovering and interpreting these meanings is the "primary goal of basic qualitative research" (Merriam, 2009, p. 24). Basic qualitative research goes beyond uncovering a participant's experiences and attempts to understand the meaning behind those experiences (Worthington, 2019).

5.3.2. Population and Subject Selections

A purposeful random sample was used to determine the participants in this study. "The logic and power of purposeful sampling lies in selecting information rich cases for in-depth study" (Patton, 2015, p. 264). By using a purposeful random sample, there is a reduction in suspicion of why certain cases were selected. As Patton (2015) describes, "purposeful random sample is its own randomness" (p. 286).

Administratively, the state of North Dakota is split into three Extension districts; west, central, and east (NDSU Extension, 2019). Of the 53 counties, 20 are in the west, 20 are in the central, and 13 are in the east.

While each Extension agent has a 4-H youth development responsibility, there is one agent in each county who serves as the lead for the 4-H program. Due to vacancies at the time the interviews were conducted, three counties did not have an agent in the lead 4-H role. For this reason, these counties were not considered for the study. North Dakota Extension agents (N=49) who held the primary responsibility for 4-H in their counties and were employed as November 9, 2018 were eligible to participate in this study.

Table 7 displays a sample of the data used to determine participants. The total population as well as the youth population as reported by North Dakota Compass in November 2018 (Wilder Research, 2018). 4-H membership numbers were collected from the statewide registration system and reflect 4-H enrollment of youth ages eight and older at the end of the 2016-2017 4-H year. These numbers were used to calculate the percentage of eligible youth in the county who participated in the 4-H program during the 2016-2017 4-H year. Counties were divided into three groups based on the percentage of youth participating, under 5%, between 5% and 10%, over 10% youth participation. From each group, three counties were chosen, one representing each program area (4-H, FCW, and ANR) as well as each district represented. Urban or rural distinction was also considered when determining participation (North Dakota Department of Commerce, 2017). This yielded a sample size of nine counties', due to one county declining to participate, an additional county was sent the recruitment email. The 4-H lead person for these counties were sent the recruitment email (Appendix C) and were invited to participate in the study.

Table 7

Youth Participation in the 4-H program in North Dakota by County

County	Total Population	Youth Population	4-H membership (8+)	Percentage	Urban/ Rural	4-H Lead
1	181,516	38,869	339	0.87	Urban	4-H
2	95,273	21,792	134	0.61	Urban	4-H
3	70,770	14,522	125	0.86	Urban	4-H
50	1,769	455	41	9.51	Rural	ANR
51	1,349	241	21	8.71	Rural	ANR
53	763	164	59	38.98	Rural	ANR

(North Dakota Department of Commerce, 2017).

Of the ten invited participants, five (50%) completed the interview process. The agents recruited for the study formed a diverse sample based on county youth population, administrative district, position held, and population category.

Table 8

Research Participants

Participant Pseudonym	Title	Years of Service
Helena	4-H Youth Development	18 months
NPA	Family & Community Wellness	25 years
Lena	4-H Youth Development	5 years
Adeline	Agriculture & Natural Resources	4 years
Ellaine	4-H & Agriculture & Natural Resources	15 years

5.3.3. Instrument Development

Mills, Durepos, and Wiebe (2010) defined interviews as "conversations with purpose and direction" (p. 495). During a semi-structured interview, participants were asked to explain how they use any assessment or evaluation in their 4-H programs. Participants were also asked to describe how they view themselves as an Extension agent and how they use and view assessment and evaluation in their programming. A semi-structured approach allows the researcher to both compare a participant's responses as well as to seek to fully understand each participant's unique experience (Mills, et al.., 2010).

Table 9

Interview Questions

Questions	Probes
Your participation in this study will only be shared between myself and Dr. Adam Marx, my advisor. For this reason, in all reports you will be referred to using an alias. Do you have a preferred alias you would like to use?	
Tell me about you as an Extension Agent	What is your role within NDSU Extension?
	How long have you been with NDSU Extension?
	What kind of work did you do prior to coming to Extension
	How would you describe your 4-H program to someone outside of Extension?
	Do you see yourself as a good Extension Agent?
How do you define assessment?	What do you think your role as an Extension Agent is when it comes to assessment?
	How do you decide if a program or event was a success?
	How do you share this success with stakeholders and/or supervisors?
	Before Common Measures, how did you assess your programs?
	If you could build your own assessment, in place of Common Measures from the ground up, what would that look like?

Table 9. *Interview questions (continued).*

Questions	Probes
I was able to pull a few of your impact statements from the NDSU website. I specifically pulled reports that you have completed around the 4-H youth development program.	What type of assessment did you use to write these? -- (If it is Common Measures) How did this report go over with your district director? How about with other stakeholders?
Now we are going to talk a little about Common Measures, when I ask questions about Common Measures, please think about any interactions you have had with Common Measures. Using only three words, how would you describe Common Measures?	Elaborate on each word chosen.
When you first found out about using Common Measures, how did you feel about the process?	Did your thoughts and feelings change over time? In what ways? What benefits do you see with using Common Measures? What drawbacks do you see with using Common Measures? Did you attend the Common Measures training in March? What impact, if any, do you think that had on how you see Common Measures?
Do you use Common Measures currently?	Why or why not? Why did you choose to do it this way? How have those in your 4-H community reacted to the new assessment? What do you plan to do with any of the data you have collected using Common Measures? - Do you have any reservations about doing this?
What do your supervisors or peers think of Common Measures?	Are you in line with these? Do you see Common Measures having any impact on the culture of assessment and evaluation in North Dakota 4-H?
What kind of community do you see among Extension Agents at NDSU?	How do you think these relationships affect your thoughts and feelings on programs and projects that are rolled out?

Table 9. *Interview questions (continued).*

Questions	Probes
	How do you use this community to benefit your program?
	Where do supervisors and administration fit in this community?
	Who do you see as the drivers of Extension? What positions do they hold?
Out of all the things we've talked about today in regard to Common Measures – or maybe some things we've missed – what should I pay most attention to? What should I think about when I read your interview?	

Table 9 lists the questions that were used as prompts to guide the interview. As is the case with guided interviews, there were instances that not all prompts were necessary to collect the needed information.

5.3.4. Data Collection

After creating the interview instruments and consent forms for participants, an application for permission to conduct the research was submitted to the Internal Review Board (IRB) at North Dakota State University. The IRB approved documents are included in Appendices A through C.

Per the approved IRB, the ten potential participants were contacted via their institutional email to request their participation in the study. Included with the email were the interview questions as well as the informed consent form. The full text of the email invitation and reminder messages is provided in Appendix C. Interviews were conducted during November 2018 at a location chosen by the participants. They were asked to sign and send the consent form back to the researcher prior to the interview taking place.

5.3.5. Data Analysis

Data collected in the study was transcribed using Rev.com, a third-party transcription service. After transcription, a series of coding was performed on the data. Coding provides a way to "translate" data that is collected into meaningful codes that are then used for data analysis (Vogt, Vogt, Gardner, & Haeffeke, 2014, p. 13). A code "capture[s] a datum's primary content and essence" (Saldaña, 2016, p. 4). Coding takes the data collected through what Saldaña (2016) describes as a cyclical process. First cycle coding was completed using In Vivo, Initial, and versus coding. After reach cycle of coding an analytical memo was written.

Second cycle coding focuses on developing a sense of "categorical, thematic, conceptual, and/or theoretical organization from your array of first cycle codes" (Saldaña, 2016, p. 234). Second cycle coding takes the codes from the first cycle and synthesizes them into more succinct themes and concepts. For the study, a second cycle coding of pattern coding was completed.

5.3.6. Controlling Trustworthiness

A debate still remains if the aim of qualitative research should be objectivity or trustworthiness (Patton, 2015). While both sides have compelling arguments, neither can be achieved perfectly. For the research study, fairness and trustworthiness will be used as indicators of non-biased results. Fairness has four defining features:

- Acknowledgment and equal representation of both sides of the case
- Equal representation of each side of the case to ensure the reader has all the facts to draw independent conclusions
- Ensuring that biases are checked during an interview and have limited involvement in the presentation of facts.

- "A relative criterion that is measured by balance rather than by isomorphism to enduring truth" (Patton, 2015, p. 725).

During the process of qualitative research, it is imperative that the researcher identify blind spots in the research and be attuned to any biases that may exist. By being aware of these biases, a researcher can look for "negative cases," a step to further improve trustworthiness (Patton, 2015, p. 685).

5.4. Findings

Of the five interviews conducted, three were done in person and two were done via Skype. Overall, agents were very open to discussing the pros and cons of Common Measures and evaluation in North Dakota.

The findings from the study are presented below as they related to each of the proposed research questions. It is important to note that several of the themes intermixed between the questions and evidence of multiple questions can be found in each of the following sections. The use of pattern coding allowed the researcher to synthesize the codes into the main ideas of the research questions. In addition to themes emerging to answer the proposed research questions, additional information regarding the culture of North Dakota Extension and the interplay between counties and agents emerged. These findings are presented in section 4.6.

5.4.1. Prior to Common Measures, how did Extension agents evaluate their 4-H program?

Overwhelmingly, agents interviewed for the study used qualitative data they collected through a variety of means to determine the success of their 4-H program. Surprisingly, none of this data was used to write any impact statements. All the impact statements pulled and discussed with the participants were focused on a single educational program, rather than the whole county 4-H program. Agents stated they could tell their programs were a success through things such as;

a waiting list, public response to the program, and publicity in the local media. This type of data collection allows an agent to determine if their program was successful and what needs to be done to improve it in the future. The fault of this form of data collection is the disconnection from the state mandated reporting requirements. “I can tell a story, but I can’t write a great impact statement,” explained Ellaine, “I can tell you what we’re doing, but telling it and putting it on paper, you lose a lot of the passion, I think.” When asked why they had not used this data to tell the story of their program in an impact statement, four of the respondents stated they did not feel it met the level of evaluation that was required from the state (3-4 on the Kirkpatrick model). The other agent stated there were different programs that rose to the top every year that she felt needed to be reported on.

Two of the agents, who both have a split appointment, stated that the impact statements they have chosen to do over the years are related to surveys that have been developed by specialists and come as part of a packaged program. One participant stated that the reason that she has never done one on 4-H is due to the lack of packaged 4-H programs with valuable surveys included. One participant stated it had been a while since she did an impact statement on just 4-H. When asked why, NPA stated that other projects and programs that provided survey tools “were far better at gathering information. I hate having to develop a survey by myself. And it’s just easier because some of the other programs that we do, there’s already [an] evaluation tool that’s provided for us.” NPA latter added that building an evaluation makes her “lose time.” This balance between a quality program and a quality evaluation tool was expressed by several participants. As Helena stated, “I feel like if we tag a survey to every single thing we do, people get bogged down with them but maybe that’s something I need to get better at is collecting data for everything I do." The struggle to meet the needs of state reports as well as continue to meet the needs of the county were very apparent.

While agents are evaluating and collecting data on their 4-H program, they are not taking the next step to report these findings and help tell the story of the 4-H program. A divide between the expectations of reporting and the tools agents are given to gather and report the data continues to grow as reporting requirements for the state and program partners evolve. This divide has caused agents to not tell the whole 4-H story.

5.4.2. What are Extension agents' perceptions of the importance of evaluation in their role as an agent?

"An assessment of learning, I guess I see that more as a test or actually seeing visual results. Whereas an evaluation is, I guess in our world, you intend to make change," said Adeline. Showing an intent to make change on the Kirkpatrick scale is at the center of how agents believe they need to report impacts. While most agents understand the importance of evaluation as an agent, not all see it as an integral part of the position. Helena described evaluation as bean counting, "I feel like there's so much bean counting and if we spend all of these hours recording all of the people we see, it's taking away from time that we could be using for our job." The formality of reports and impact statements is where agents tend to start seeing the evaluation as more of a requirement and less of a benefit to their program.

The most common way data is collected by agents for reporting purposes is a survey after the program. In some instances, these surveys are done before participants leave the program, and in others, are done as a follow up several months after completion of the program. Data collected this way can be used to develop statistical reports in support of the program. Yet, often this data isn't valuable to an agent to evolve the program. "I feel like a part of us needs to appease the state, and the other part of us needs to appease our county. I would rather know how the program is doing," said Helena who further elaborated,

> what I can do better for the people in my county. And what programs they're interested in and want to learn more about. I guess, I am more concerned with what they think and feel, but I think sometimes the state has certain things they want us to find out, and then I don't want to say push their agenda, but.

Agents often see the evaluation tool to meet the reporting guidelines as an afterthought, or something that is pushed aside due to lack of time. All the participants in the study stated that they gather data to determine the success of a program through informal qualitative means. This data is used to help refine programs for the future and adjust as needed but isn't used to report on the successes of the program through formal means of an impact statement.

When asked for her thoughts on Common Measures and the importance of assessment Ellaine said;

> I think it [assessment] is important, but I think it has to be done so it's effective for the communities we serve. When I started there was a CYFAR grant, and the pre and post-test were both 35 pages long. It was rather disastrous because I don't think the information they received back was accurate because people were burned out. A lot of the questions weren't relevant to the population. They got data, but I don't think it was accurate. I think if you give data, it needs to be accurate.

Similar to Ellaine, other participants see the importance of the data being specific to the county they are serving to help tell the story. Collecting data the 'traditional' way through the administration of a survey does not work for all populations that agents in North Dakota serve, "literacy skills are not always the best by some of our older members and parents of our youth…so we do verbal, and I gather data that way" stated Ellaine. It is important to note, that Ellaine gathers all her data this way. If a survey is requested for a specific program, she will use oral assessment to collect the data that is necessary to meet the reporting requirements.

Extension helps shape the narrative of North Dakota, especially in the rural areas of the state. Three of the participants stated that they feel evaluations provide them with the information they need to better tell the story of the 4-H program. As Adeline who is from a rural county stated, rural counties are often over looked in the big picture of the program. Data from surveys, such as Common Measures, gives her the resources to be able to say, "hey, we are here. And we have some amazing youth in the county, and they have the potential to do whatever they set their mind to." The narrative of the importance of rural youth and families is not lost on agents, and 4-H allows Extension to have an increased connection with these important stakeholders. Leana described 4-H as the "gateway to youth."

5.4.3. Do Extension agents view new programs and/or procedures, such as Common Measures, as intrinsic or extrinsic?

Overwhelmingly, agents felt that Common Measures was extrinsically motivated. "I didn't want to mess up their data" said Lena when discussing how she plans to use the Common Measures data to tell the story of her county program.

> I have pulled from blocks of questions in the past, when I've tried to evaluate things, but I kind of feel like what you just said, it doesn't get at the beef of what the stakeholders might want to know

continued Lena. From a broader perspective, agents felt that reporting and evaluation was mandatory and done to appease the state. Interestingly, after further conversations, agents reported that they do evaluation on a regular basis of their program, however, it is informal in nature and does not yield the necessary data that is required for reporting. However, every agent interviewed said they use alternative forms of data collection (school data, enrollment numbers, focus groups) to gather data about the health of their program and what could be done better. This data is often not viewed as data that is good enough to report findings to the state. When asked about evaluation, none of the agents mentioned this type of data collection, even though all of them are collecting data in these non-traditional ways. While participants see the value in non-traditional data collection, they lack the understanding and skill to transfer that to the necessary reports for the state. All of those interviewed stated that qualitative information helps them better address the needs of their county.

Much of the data that agents are informally collecting helps paint the picture and tell the story of their county 4-H program. Ellaine shared a success story of a youth in her county who was the first in her family to go to school beyond high school who recently graduated with her master's degree. As Ellaine put it, "those are the kind of impact statements I think we need to be getting out there." The data collected for this impact was not done on a survey, it was not an impact that was achieved in a year, or even five, it took years of programming and development of the young person to see the results. These are the stories that agents tell, however these same stories are missing from the state required impact statements.

A few agents reported that their clubs and members understood the importance of doing evaluations, and, while they didn't like it, were willing to complete the surveys if it helped the 4-H program. NPA described Common Measures as "cumbersome" and "missing the mark" on several items. "That's Common Measures, it's a long survey, it's a long evaluation tool. Do these young people, do they have the patience with their attention space to sit through it and finish it? Are they just checking boxes?" stated NPA.

On the other side of the spectrum, when Helena was asked if she had a better understanding of the questions and the data they provided if she would use Common Measures she said, “I feel like if we tag a survey to everything single thing we do, people get bogged down with them.” As a follow-up Helena was asked if she felt Common Measures could replace some of the surveys she is currently using, she said

> I wasn’t super impressed with the questions that they had listed or if it would pertain to the program that I would do […] I think it’s great and it’d be a good indication […] that kids do like 4-H but I don’t think it explains why they like 4-H so much.

Overwhelmingly, agents lack the knowledge and skills necessary to take the information they collect in non-traditional ways and put together a report that tells the story of their 4-H program. Because of the lack of connection between non-traditional data and reports, agents see evaluation as checking a box, completing a requirement from the state, and extrinsically motivated.

5.4.4. What training or professional development is needed in North Dakota to improve evaluation of the 4-H program?

To create a better connection between the data agents are collecting to improve their program and the state mandated reports, agents across the state would benefits from additional training and resources. Training in the areas of qualitative data collection, analysis, and reporting, background of questions asked on Common Measures, developing an evaluation tool from scratch, and support after trainings are delivered were common themes throughout the interviews.

Agents are collecting qualitative data to personally gauge the success of the program. However, they are not using that data to report on the impact of their program. The research indicated that agents lack the knowledge necessary to use data collected this way to report program successes. While Common Measures is quantitative in nature, agents feel that a story is needed to accompany the numbers. Ellaine explained,

> if one of my kids [4-H members] doesn't understand a science project that we're doing, I want to know why they're not understanding it. I just don't want to say, 'do you like science?' because to me, there's no depth to that. I like things in depth. If I'm going to do an evaluation, I want to get some real hard-core data.

Agents have the quantitative data and stories however are not equipped with the training and tools necessary to make the connection to the reporting required by the state.

Along the lines of meeting the reporting requirements, agents felt that more training was needed on how to develop a survey tool that reaches this goal as well. One agent stated it would be very beneficial if a specialist "took a program and then used Common Measures data to write and impact statement." This approach would show the agent "this is how you can use the data effectively." Several participants stressed the disconnection between Common Measures and the state required impact statements. Agents need better direction on how to connect the two. Additionally, district directors and Extension leadership would benefit from training on Common Measures. Ensuring that leadership and agents are functioning with the same knowledge will create a better understanding of 4-H program evaluation on all levels of the organization.

Participants did not feel as if they had enough knowledge on the background of the questions on Common Measures to fully understand the entire project. Even though a Helena attended the state-wide training on Common Measures, she commented that she was "not super well-versed in it [Common Measures]." Specifically, questions such as, "how many hours do you typically spend on 4-H activities each week"(p.9) which are explained by research, do not resonate with agents (National 4-H Council, 2018). By providing training and background information more specific to the tool, may help agents understand the rationale behind the questions asked.

Both agents that attended the training sponsored by the state, stated that upon returning to their counties they felt like they did not have the knowledge needed to successfully implement the survey successfully in their counties. Better tools and support once agents have left a training and are back in their counties is imperative to program success. Additionally, providing tools and resources to help agents get through different situations that may arise is necessary. As Ellaine stated, “make a program fit the people you serve.” Knowing how, and when you can manipulate the tool to get better data that not only gets the data that is needed need for the state and national reports, but also provides the local program with information that will contribute to the progress and success of the program. Specifically, it was expressed that working with diverse populations such as kids not enrolled in a 4-H club (but participating in Extension youth development activities), and younger youth (Cloverbud age) were not addressed by the tool and create barriers that agents must overcome to be successful in collecting data.

Agents expressed concerns with needing to find a balance between delivering high quality content and having a high-quality evaluation. Through a wider lens, agents expressed the need for more planning and collaboration sessions that help achieve the goal of the evaluation being a part of the planning process rather than an afterthought. States that have a dedicated evaluation specialist were brought up as a way that would provide beneficial support to agents across the state. Adeline explained,

> Don't just say, "Well, here it is." Teach me how do I know what to ask? Or what's the data we're looking for? I can think in my head, like, "Okay, I know I want this." But then,

> jeepers. I get the results back and I'm like, "What the heck? This is nothing. I should have asked this." Just being able to help think through that process.

Agents need help not only forming the evaluation, but also analysis the data so it is valuable for required state reports as well as program development and improvement.

5.4.5. How does an agent's length of service with Extension, position, education level, and county composition influence differing viewpoints of evaluation?

Overall, an agent's length of service had a positive influence on their views of evaluation. Helena, who has been with Extension for the shortest period had very negative thoughts and feelings about any formal evaluation tool. Formal education in Extension did not seem to coincide with any of the participant's thoughts on Common Measures. It was often participants with no formal educational background in education or Extension that saw the most value in a standardized evaluation tool.

Lena, who has five years of experience in Extension stated, "I think it [evaluation] is a really important part of being an Extension Agent that I didn't truly appreciate or fully understand until more recently." Lena stated throughout her interview that much of the data she uses to define a successful program comes from conversations with people in and around her county. Lena comes from a rural county

NPA, whom has 25 years of experience stressed the importance of an evaluation tool, like Common Measures, on the 4-H program across the state. She was supportive of the idea of the tool, however, did not feel that Common Measures answers the questions that needed to be answered to achieve the result that was trying to be achieved. "I was glad that we finally had a formal tool, when I looked at the tool, you know, some of the questions just didn't quite seem to hit the mark."

All participants in this study are considered the 4-H lead for the county they work in. However, those that hold a split role often did not focus their impact statements on the 4-H program. NPA and Lena both stated that the ease of using a premade evaluation tool that is available with packaged programs helps them determine the program they will formally report on. Due to the nature of the 4-H program, especially in North Dakota, packaged programs are very limited, and thus, 4-H agents are not able to use a pre-made evaluation tool and must develop one of their own.

Due to the small sample size and limited areas considered urban in North Dakota, no conclusions could be drawn on the difference between rural and urban counties as it related to evaluation.

5.4.6. Additional Themes

Throughout the research, themes not related to the research questions emerged in the data that are worth noting. Every participant in the study noted a county vs. county competition that exists in North Dakota. Several times, this was brought up in a negative way and as something that needs to be addressed in the state for Extension to continue to move forward. Adeline referenced a former document that was often updated and circulated within Extension that color coded counties based on membership changes over the last 5-10 years.

> It was tough when Dean would make that freaking map of North Dakota. To look at my county being red… Red. Red, red, red. Why is it the amount of kids that matters? Why isn't it the experiences that we're giving them? That can be different, it is different within every county.

The red Adeline is referring to noted counties that had not seen a membership increase over a certain amount. Adeline continued by saying that she is hopeful Common Measures will help change the narrative from county vs. county to more experiences based. While a bit of competition is healthy, it should be noted that several participants stated they feel the competition often hinders program development across county lines as well as program areas.

As Lena said, “I think there is a disconnect about what 4-H is, even internally.” Agents with a split roll described this more of a balancing act, a constant push and pull between 4-H and their other program area. Even though every agent is required to have a role in 4-H youth development, counties with multiple agents thinks this looks different if they are not the 4-H lead for the county. Lena said that NDSU Extension could benefit from more cross-programmatic programming,

> there already has been a lot of focus on cross-programmatic planning and I think the more from the top we are encouraged to use our specialists in other disciplines, partner, I really think the more we partner, even as counties, on things, we build these relationships that open us up to new resources. And I kind of think a lot of that is some counties already do it so well, so strongly. They help each other out all the time and other counties have their distinct disciplines and so I guess the more it's encouraged, that community will become stronger.

Cross-programmatic planning would hopefully alleviate some of the county competition that agents reported during

5.4.7. Findings Conclusions

Overall, the agents that participated in the study were initially hopeful for Common Measures and the data it could provide stakeholders at all levels of the organization. However, after the initial implementation of the survey, agents feel it could use some additional refinement. In addition to refinement it is important to note the needs that arose out of the research for additional training and resources. Given the right tools, agents are willing and able to use evaluation and assessment to address reporting requirements as well as local needs.

5.5. Conclusions and Recommendations

The growing pressure on agents to deliver impactful reports of the work being done in the county often has agents in a constant state of push and pull between doing an impactful program and developing an impactful evaluation tool. A further push can be seen for the impacts to be transformative in nature. While transformation does occur through Extension youth programming, it is often not seen in a year nor is the change best captured with a formal evaluation, often a survey, tool. Common Measures was a way to help agents feel the need for an evaluation tool to help tell the story of their 4-H program, however, the tool feel short and did not meet the needs of agents for a variety of reasons.

While agents would like to see a more specific and program focused evaluation, Common Measures was created to have a common tool to measures 4-H programs across the nation. With this in mind, most of the recommendations are based on additional training for agents and specialists at NDSU. The complete list of recommendations is as follows:

- Further training for staff on how to use the Common Measures tool to better tell their story and meet the reporting requirements for the state. With the data already collected agents are set to use it to highlight their 4-H program successes. However, the tool does not allow flexibility for agents to target to their specific needs and thus is not seen by agents as having any value.
- NDSU Extension needs to establish relevance of Common Measures and other evaluation tools with agents. As Patton (1997) discusses, having a shared definition and commitment to the process contributes to the success of any evaluation method.

The process toward this shared definition and commitment should be obtained through working committees of agents and specialists. As agents reported during the interviews, agents are the drivers of the organization. Better utilization of agents in these roles will improve success.

- Provide training on using qualitative data to help tell the story of the 4-H program. Agents would benefit from learning how to collect the data as well as how to analyze and present it in such a way that it is meaningful to partners, stakeholders, and meets the requirements of the state. Many agents are collecting qualitative data, however are not using it in any meaningful report. As Worker (2019) illustrated in his use of qualitative data to assess a 4-H program, clearly defining the parameters is imperative to a successful study. Agents need training and support to develop these skills in order to effectively collect and report data to tell the story of their programs. An additional need of providing survey tools that can be used across the state, including in Indian Country, can also be addressed through the use of qualitative data collection.
- Agents reported that they see agents as the drivers in Extension. Similar to the process of rolling out Common Measures, committees of agents and specialists should be used to roll out state wide initiatives. Not with standing, it would also be valuable to agents to have the skills and tools necessary to include the populations being evaluated in survey development. As Martinez, et. Al (2018) points out, including Native Americans in the development and execution of data collection in Indian country would also benefit agents working with the Native population.

The success of the North Dakota 4-H program is only as strong as the story. If North Dakota Extension continues to share these stories through reports and impact statements, it is vital that agents, are given the tools, training, and resources necessary to not only delivery high quality programs, but also evaluate them in a way that is meaningful to their local program as well as the state program.

As Common Measures continues to be used by agents and specialists in North Dakota a better understanding of how Common Measures is viewed by not only agents but also specialists and mangers is necessary. Given the newness of Common Measures, additional studies focused on the use and development of the instrument could prove beneficial as National 4-H move forward with a nationwide survey tool. Recommendations for additional studies include:

- A similar qualitative study with specialists and members of management would not only give a better understanding of additional trainings needed for those areas but would provide insight into the disconnect between agents and supervisor expectations for programing, evaluation, and reporting.
- A better understanding of how youth and adult volunteers view the tool would also aid in the development of additional trainings and tools related to Common Measures. As those who the tool is geared toward, youth can provide valuable insight into some of the questions that agents felt were not relevant to youth.
- A qualitative study across the nation of educators using the tool would provide a better picture of how agents are using the tool and potential practices that could be implemented in North Dakota

REFERENCES

Bailey, S. J., & Deen, M. Y. (2002). A Framework for Introducing Program Evaluation to Extension Faulty and Staff. *Journal of Extension*.

Braverman, M. T., & Arnold, M. E. (2008). An Evaluator's Balancing Act: Making Decisions About Methodolocial Rigor. *New Directions for Evaluation*, 71-86. doi:https://doi.org/10.1002/ev.277

Committee on the Future of the Colleges of Agriculture in the Land Grant University System. (1995). *Colleges of Agriculture at the Land Grant Universities.* Washington, D.C.: National Academy Press.

Donaldson, J. L. (2019, July 6). *Extension Evaluation and Staff Development.* Retrieved from The University of Tennessee : https://extension.tennessee.edu/eesd/Documents/SUPER/HelpFiles/PEN_UserGuide.pdf

Fiester, L. M., Simpkins, S. D., & Bouffard, S. M. (2005). Present and accounted for: Measuring attendance in out-of-school-time programs. *New Directions for Youth Development*, 91-107. doi:https://doi.org/10.1002/yd.109

Flage, L. (2019, February 7). *About Us*. Retrieved from Leadership and Civic Engagement: https://www.ag.ndsu.edu/lead/about-us

Fournier, D. M. (2005). *Encyclopedia of evaluation.* Thousand Oaks: Sage.

Franz, N. K., & Townson, L. (2008). The nature of complex organizations: The case of cooperative Extension. *New Directions for Evaluation*, 5-14. doi: https://doi.org/10.1002/ev.272

Franz, N., Arnold, M., & Baughman, S. (2014). The Role of Evaluation in Determining the Public Value of Extension. *Journal of Extension*.

Great School Partnership. (2014). *Assessment*. Retrieved from The Glossery of Education Reform: https://www.edglossary.org/assessment/

Hawley, L. R. (2017). *4-H common measures 2.0: General user's guide and protocol.* Lincoln: Nebraska Academy for Methodology, Analytics & Psychometrics.

Hill, P., & Seger, J. (2018). *We've Tried That Before: 500 Years of Extension Wisdom.*

Jayaratne, J. (2013). *Introduction to Evaluation Tools*. Retrieved from NC Cooperative Extension: https://evaluation.ces.ncsu.edu/evaluation-introduction-to-evaluation-tools/

Kirkpatrick Partners. (2018, December 16). *The Kirkpatrick Model*. Retrieved from Kirkpatrick Partners: https://www.kirkpatrickpartners.com/Our-Philosophy/The-Kirkpatrick-Model

Kirkpatrick, D. (1996, January). Great Ideas Revisited. *Training and Development*, pp. 54-59.

Kushner, J., Crave, M., & Jones, L. (2017). *Assessing System Gaps in Extension Education Evaluation Resources.* Madison: University of Wisconsin- Extension.

Lau, G., Netherland, N. H., & Haywood, M. L. (2003). Collaborating on Evaluation for Youth Development. *New Directions for Evaluation*, 47-59. doi:https://doi.org/10.1002/ev.84

Lerner, R. M., & Lerner, J. V. (2013). *The Positive Development of Youth: Comprehensive Findings from the 4-H Study of Positive Youth Development.* Chevy Chase: National 4-H Council.

Lewis, K. M., Horrillo, S. J., Widaman, K., Worker, S. M., & Trzesniewski, K. (2015). National 4-H Common Measures: Initial Evaluation from California 4-H. *Journal of Extension*.

Martinez, A., Running Wolf, P., BigFoot, D. S., Randall, C., & Villegas, M. (2018). The Process of Becoming: A Roadmap to Evaluation in Indian Country. *New Directions for Evaluation*, 33-45. doi:https://doi.org/10.1002/ev.20333

Martz, J., Mincemoyer, C., & McNeely, N. N. (2016). *Essential Elements of 4-H youth development programs*. Chevy Chase, MA: National 4-H Council. Retrieved from https://fyi.extension.wisc.edu/wi4hstem/files/2015/02/full-training-curriculum-and-appendices.pdf

Mashburn, D. (2019, July). *AIMS Help: Reporting and Impact*. Retrieved from Cooperative Extension System: https://www.uaex.edu/support-units/program-staff-development/program-evaluation/impact.aspx

Merriam, S. (2009). *Qualitative research: A guide to design and implementation.* San Francisco, CA: Jossey-Bass.

Mills, A. J., Durepos, G., & Wiebe, E. (2010). *Encyclopedia of Case Study Research.* Thousand Oaks: Sage Publications.

National 4-H. (2010, Novemeber 1). *4-H Science Logic Model.* Retrieved from https://4-h.org/wp-content/uploads/2016/02/4-H-Science-Logic-Model.pdf

National 4-H Council. (2015). *Common Measures FAQ.* Chevy Chase, MD: National 4-H Council.

National 4-H Council. (2018). *4-H Common Measures 2.0 Reference Table.* Chevy Chase, MD: National 4-H Council.

National 4-H Council. (2019, May 12). *4-H History*. Retrieved from 4-H: https://4-h.org/about/history/

National 4-H Council. (2019). *What is 4-H?* Retrieved from 4-H: https://4-h.org/about/what-is-4-h/

National 4-H History Preservation Team. (2019, June 21). *4-H Motto, Creed, and Pledge*. Retrieved from National 4-H History Preservation Program: https://4-hhistorypreservation.com/History/M-C-P/

National Institute of Food and Agriculture. (2019a, January 1). *Cooperative Extension History*. Retrieved from National Institute of Food and Agriculture: https://nifa.usda.gov/cooperative-extension-history

National Institute of Food and Agriculture. (2019b). *1890 Land-Grant Institutions Programs*. Retrieved from USDA Programs: https://nifa.usda.gov/program/1890-land-grant-institutions-programs

National Institute of Food and Agriculture. (2019c). *Cooperative Extension System*. Retrieved from National Institute of Food and Agriculture: https://nifa.usda.gov/cooperative-extension-system

National Institute of Food and Agriculture. (2019d, January 1). *Extension*. Retrieved from National Institute of Food and Agriculture: https://nifa.usda.gov/extension

Native American Center for Excellence. (2019, Sept 2). *Steps for Conducting Research And evaluation In Native Communities.* Retrieved from Native American Center for Excellence: https://www.samhsa.gov/sites/default/files/nace-steps-conducting-research-evaluation-native-communities.pdf

NDSU. (2008a, July 11). *Evaluation*. Retrieved from NDSU Extension: https://www.ag.ndsu.edu/evaluation

NDSU Archives. (2019, April 25). *North Dakota State University Archives*. Retrieved from NDSU Historical FActs: https://library.ndsu.edu/ndsuarchives/ndsu-historical-facts

NDSU Extension. (2008, August 1). *Impact Statements*. Retrieved from North Dakota State University: https://www.ag.ndsu.edu/impactstatements/impact-statement-guidelines/impact-report-faqs

NDSU Extension. (2019, June 20). *Extension Agent, Agriculture and Natural Resources – Horticulture/Cass County.* Retrieved from https://www.ag.ndsu.edu/careers/current-job-openings/job-openings-pdfs/copy_of_ExtensionAgentANRHorticultureCassCounty.pdf

NDSU Extension. (2019, April 1). *NDSU Extension*. Retrieved from Directory: https://www.ag.ndsu.edu/extension/directory/state-map-with-districts-pdf

NDSU Extension. (2019). *Who We Are and What We Do.* Fargo: NDSU Extension.

NDSU Extension Service. (2017, June). *North Dakota State University Extension Service.* Retrieved from Evaluation: https://www.ag.ndsu.edu/evaluation/documents/county_narrative.pdf

North Dakota Department of Commerce. (2017, December). *Growing ND by the numbers.* Retrieved from https://www.commerce.nd.gov/uploads/26/CensusNewsletterDec2017.pdf

North Dakota Indian Affairs Commission. (2010). *Tribal Nations*. Retrieved from North Dakota Indian Affairs Commission: http://indianaffairs.nd.gov/tribal-nations/

Patton. (2015). *Qualitative Reserach & Evaluation Methods.* Thousand Oaks, CA: Sage.

Patton, M. Q. (1997). *Utilization-Focused Evaluation* (Vol. 3rd ed.). Thousand Oaks, CA: Sage Publications.

Patton, M. Q. (2008a). *Utilization-Focused Evaluation.*

Patton, M. Q. (2008b). sup wit eval ext? *New Directions for Evaluation*, 111-115. doi:https://doi.org/10.1002/ev.279

Payne, P. B., & McDonald, D. A. (2012). Using Common Evaluation Instruments Across Multi-State Community Programs: A Pilot Study. *Journal of Extension, 50*(4).

Poling, R. L. (2005, February). *University of Arkansas.* Retrieved from Cooperative Extension Service: https://ag.purdue.edu/extension/pdehs/Documents/EVALPROGRev022005.pdf

Raidl, M., Johnson, S., Gardiner, K., Denham, M., Spain, K., Lanting, R., . . . Barron, K. (2004). Use Retrospective Surveys to Obtain Complete Data Sets and Measure Impact in Extension Programs. *Journal of Extension*.

Roucan-Kane, M. (2008). Key Facts and Key Resources for Program Evaluation. *Journal of Extension*.

Saldaña, J. (2016). *The coding manual for qualitative researchers.* Thousand Oaks, CA: Sage.

Sanders, H. (1966). *The Cooperative Extension Service.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Stevahn, L., King, J. A., Ghere, G., & Jane, M. (2005). Establishing Essential Competencies for Program Evaluators. *American Journal of Evaluation, 26*(1), 43-59. doi:https://doi.org/10.1177/1098214004273180

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Technigues and procedures for developing grounded theory.* Thousand Oaks, CA: Sage.

Stup, R. (2003). Program Evaluation: Use It To Demostrate Value to Potential Clients. *Journal of Extension*.

Texas A&M AgriLife Extension Service. (2019, July 7). *Our Evaluation Perspective*. Retrieved from Organizational Development: https://agrilife.org/od/evaluation/about-evaluation/

Vogt, W. P., Vogt, E. R., Gardner, D. C., & Haeffeke, L. M. (2014). *Selecting the right analyses for your data: Quantitative, qualitative, and mixed methods.* New York, NY: Guilford.

Wilder Research. (2018, November). *Population under age 18 by North Dakota Counties 2000-2018*. Retrieved from North Dakota Compass: https://www.ndcompass.org/

Worker, S. M. (2019). Development of an Artifact-Based Evaluation Framework for Assessing 4-H Learner Outcomes. *Journal of Extension*.

Worthington, M. (2019, July 6). *Differences between Phenomenological Research and a Basic Qualitative Research Design.* Retrieved from http://a1149861.sites.myregisteredsite.com/DifferencesBetweenPhenomenologicalResearchAndBasicQualitativeResearchDesign.pdf

APPENDIX A. IRB APROVAL LETTER

NDSU NORTH DAKOTA STATE UNIVERSITY

September 14, 2018

Dr. Adam Marx
School of Education

Re: IRB Determination of Exempt Human Subjects Research:
Protocol #HE19051, "Common Measure adoption in North Dakota 4-H"

Co-investigator(s) and research team: Amelia Doll
Date of Exempt Determination: 9/14/2018 Expiration Date: 9/13/2021
Study site(s): varied
Sponsor: National 4-H Council (FAR0028609)

The above referenced human subjects research project has been certified as exempt (category #2b) in accordance with federal regulations (Code of Federal Regulations, Title 45, Part 46, Protection of Human Subjects). This determination is based on the original protocol submission (received 9/12/2018).

Please also note the following:
• If you wish to continue the research after the expiration, submit a request for recertification several weeks prior to the expiration.
• The study must be conducted as described in the approved protocol. Changes to this protocol must be approved prior to initiating, unless the changes are necessary to eliminate an immediate hazard to subjects.
• Notify the IRB promptly of any adverse events, complaints, or unanticipated problems involving risks to subjects or others related to this project.
• Report any significant new findings that may affect the risks and benefits to the participants and the IRB.

Research records may be subject to a random or directed audit at any time to verify compliance with IRB standard operating procedures.

Thank you for your cooperation with NDSU IRB procedures. Best wishes for a successful study.
Sincerely,

Kristy Shirley

Kristy Shirley, CIP, Research Compliance Administrator

For more information regarding IRB Office submissions and guidelines, please consult http://www.ndsu.edu/research/integrity_compliance/irb/. This Institution has an approved FederalWide Assurance with the Department of Health and Human Services: FWA00002439.

INSTITUTIONAL REVIEW BOARD
NDSU Dept 4000 | PO Box 6050 | Fargo ND 58108-6050 | 701.231.8995 | Fax 701.231.8098 | ndsu.edu/irb

Shipping address: Research 1, 1735 NDSU Research Park Drive, Fargo ND 58102

APPENDIX B. INFORMED CONSENT

Common Measures: A case study of assessments in North Dakota State University Extension

Amelia Doll and Dr. Adam Marx
North Dakota State University

Why am I being asked to take part in this research study? Because you are a current employee of NDSU Extension, we are interested in learning about your experiences with Common Measures and the use of assessment in your 4-H program.

What is the reason for doing this study? The purpose of this study is to investigate the experiences Extension Agents have had with Common Measures. The goal of this research is to expand understanding of how Extension Agents feel about new projects or programs are rolled out and what influences their reactions.

Where is this study going to take place, how long will it take, and what will I be asked to do? This study involves you participating in a face to face interview (approximately 1 hour) that will be recorded and transcribed. During our interview, we will discuss some of your impact statements that are posted on the NDSU website. You will not be asked to provide these. The interview will also include the use of the social ecological model tool that we will use to facilitate discussion.

What are the potential risks and discomforts? What are the expected benefits of the study? Any information you share with us will be kept confidential. As such, we don't anticipate any adverse effects from participating in this study. While you may not get a direct benefit from being in this research study, we hope to use the results to enhance future Extension work.

Do I have to take part in the study? What are the alternatives to being in this research study? Your participation in this research is your choice. If you choose not to participate in this study, we will not proceed with the interview and you can decline to sign the consent form. If you *do* decide to participate in the study, you may change your mind and withdraw at any time without penalty.

Who will see the information that I give? We will keep private all data that may identify you. Your information will be combined with information from other people taking part in the study. When we write about the study, we will write about the combined information that we have gathered. We may publish the results of the study; however, we will keep your name and other identifying information private. All records will be kept on password protected computers. At the conclusion of the study all identifiers will be removed from the data files.

Questions? If you have any questions about this study, you can contact the principal investigator, Dr. Adam Marx at Adam.Marx@ndsu.edu or 701.231.7439.

What are my rights as a research participant? The role of the Human Research Protection Program is to see that your rights are protected in this research; more information about your rights can be found at: www.ndsu.edu/irb. You have rights as a participant in research. If you have questions about your rights, or complaints about this research, you may talk to the research or contact the NDSU Human Research Protection Program at 701.231.8995 (or toll-free 1.855.800.6717), by email at ndsu.irb@ndsu.edu or via mail at NDSU HRPPP Office, NDSU Dept. 4000, PO BOX 6050, Fargo, ND 58108-6050.

Documentation of Informed Consent. You are freely making a decision whether to be in this research study. Signing this form means that: (*feel free to keep a copy for your records as well*)

1. You have read and understood this consent form
2. You have had your questions answered, and
3. You have decided to be in the study.

______________________	______________________	____________
Your Signature	Your Printed Name	Date
______________________	______________________	____________
Researcher Signature	Researcher Printed Name	Date

APPENDIX C. INTRODUCTION EMAIL

Appendix B:
Interview Request Email

Hello,
I am a graduate student working with Dr. Adam Marx to explore the experiences Extension Agents have had with Common Measures. The purpose of this study is to investigate the experiences Extension Agents have had with Common Measures. The goal of this research is to expand understanding of how Extension Agents feel about new projects or programs are rolled out and what influences their reactions. To achieve this, I am conducting interviews with Extension Agents in North Dakota to learn about their county program and their experiences with Common Measures. I am writing to see if you, as a NDSU Extension Agent, would be willing to participate in an interview about your experiences with Common Measures. I anticipate the interview will take about 1 hour.

For purposes of this research, I need to record the discussion and transcribe the recording – but I will leave your name (and other identifying information) out the final report. Your interview will be combined with other agents from the state in my final presentation. Your participation in these interviews will only be known to Dr. Marx and myself.

To give you an idea of the type of questions I will be asking, I have attached my interview guide for your review. As you can see, we will be using a few different tools within our time together. You will not be asked to provide any of this information, I will have all of it with me.

There are several ways we can complete the interview; I will travel to you, or they can be conducted via phone or Skype. Also attached is the informed consent form. If you are willing to participate, please review the attached document and sign and return to me. If you are willing to participate please let me know which of the following times would work for you:
[List available times]
Thank you for your consideration.
[Signature]

Follow- up email

Hi [participant name],
I am following up on the below email. I am wondering if you would be willing to participate in this research study. There is no right or wrong answer, and your identity will be protected.
Thank you for your consideration,
Amelia