FACTORS INFLUENCING PROGRAM IMPACT EVALUATION IN COOPERATIVE EXTENSION

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

By

Jan Ardis Flack

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Program:
Education
Option:
Occupational and Adult Education

October 2019

Fargo, North Dakota

North Dakota State University
Graduate School

Title

FACTORS INFLUENCING PROGRAM IMPACT EVALUATION IN COOPERATIVE EXTENSION

By

Jan Ardis Flack

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

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Dr. Claudette Peterson
Chair

Dr. Brent Hill

Dr. Ann Braaten

Dr. Carrie Johnson

Approved:

October 31, 2019
Date

Dr. Chris Ray
Department Chair

ABSTRACT

Impact evaluation (IE) has become a major focus of Extension program evaluation during the past three decades, yet Extension professionals continue to struggle to produce well-documented evidence of program impacts (Lamm, 2011; Workman, 2010).

This research was primarily a quantitative, non-experimental correlational study of factors associated with Extension educators' IE behaviors and practices in the North Central Region of the Cooperative Extension service. The study employed a comprehensive, systems approach to explore interrelated individual and organizational factors that affect IE in Extension.

A major feature of this research design was the use Structural Equation Modeling (SEM) statistical analysis methodology. The use of SEM to organize and implement this study facilitated simultaneous exploration of many constructs theorized to be involved in IE behaviors in the context of Extension educational program evaluation. Specification of a theoretical, conceptual model to be used to frame the potential relationships among the many constructs and factors was necessary prior to data collection in order to be able to use SEM for data analysis. Data were analyzed using SEM path analysis to determine relationships among the factors.

Significant findings included the identification of the most influential factors on the dependent variable of actual IE behaviors. These were: competency by perceived skill level, behavioral intention, number of roles in IE, education level, and attitude. Factors exhibiting noteworthy influence on factors other than the dependent variable within the model included training, culture, and proportion of teamwork.

Qualitative data themes most frequently mentioned by participants included: the need to know how to measure change; a need for training in general; a need for better planning of programs to achieve impact; conflicting priorities, lack of time and timing/coordination

concerns; changing expectations (“moving target”) regarding IE practices and goals; and diverseness in contexts, competencies, and schedules.

Study findings identified a number of influential factors not previously cited in literature, including number of roles in IE; proportion of teamwork employed in IE; a lack of understanding of how to measure change; and issues of diverseness regarding competencies, context (including stakeholder expectations), and scheduling within the Extension organization.

ACKNOWLEDGEMENTS

There are so many people that contributed to the successful completion of this dissertation, to whom I am very grateful. My heartfelt thanks goes out to the following members of my doctoral committee who worked with me to navigate the doctoral research process: to Dr. Claudette Peterson, my adviser, for her continuous guidance and support during the research process; she is an excellent professional and a very caring colleague; to Dr. Brent Hill, who worked closely with me to help me learn structural equation modeling (SEM) and helped me trust my instincts as I used SEM to analyze my data; to Dr. Ann Braaten, for her expertise in research which helped challenge and support my efforts; and to Dr. Carrie Johnson, whose expertise in evaluation and in the Extension Service contributed valuable insights.

Much thanks, also, to these former committee members who assisted in research design, recruitment of participants, and helped me to focus and communicate my research goals: Dr. Myron Eighmy, Dr. Deb Gebeke, Dr. Elizabeth Roumell, and Dr. Chris Ray. Thanks, also, to the many professors who guided me in this journey of personal growth and professional enrichment.

Appreciation goes out to my family and friends who encouraged and supported me by listening, providing feedback, and being patient during my doctoral journey.

Most important, special thanks to my husband, John, who supported me throughout my coursework, through the many challenges during the research process, and who always knew I could accomplish my goal of completing my Ph.D.

This has been an exhilarating, transformative process for me and I am thankful for my own strength and determination to meet the many challenges encountered. These qualities were inspired in me by my parents, Marilyn and Bob Zielke, for whom I am deeply grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Statement of the Problem

In an era of increased accountability and possible funding cuts, Extension must demonstrate relevance, impact, and return-on-investment to funders and other stakeholders regarding their educational programs (Kluchinkski, D., 2014; Decker, 1990; Kelsey, 2008; Rennekemp & Engle, 2008). To do so requires Extension educators to acknowledge that impact assessment must be considered an essential component of their educational mission (Arnold, 2015; Vengrin, 2016; Wells-Marshall, 2012).

Impact evaluation of Extension programs must be conducted with rigor and expertise not only to meet funder requirements but, also, to show the potential public value and impact to society (Kalambokidis, 2017). Impact is defined as "the reportable, quantifiable difference or potential difference a program makes in the lives of real people. It shows a sustainable societal, environmental, and/or economic change," (eXtension Impact Online course, 2017).

Evaluation of educational programming in Extension is often mandatory given that most financial support for the organization is public funding from national, state, and local sources. Increased pressure for public accountability and federal reporting compliance has prompted Extension administrators to ask Extension educators at all levels of the organization to engage in their evaluation responsibilities in a more scholarly and purposeful way (Arnold, 2015; Franz, Arnold, & Baughman, 2014). Impact evaluation has become a major focus of Extension program evaluation during the past three decades, yet Extension professionals continue to struggle to produce well-documented evidence of program impacts (Lamm, 2011; Workman, 2010).

Program Planning and Impact Evaluation in Cooperative Extension

Extension utilizes an objective-driven approach to program planning and, in turn, to evaluation to assess the impacts of its educational programs (Arnold, 2015; Wells-Marshall, 2012; Vengrin, 2016). Objectives provide guidance for program instructional strategies as well as a basis for communicating outcome expectations to participants and to program stakeholders. Two approaches to program evaluation and impact assessment have traditionally been taken. First, and the most common, is planning an evaluation AFTER a program is planned to satisfy demands of internal or external stakeholders (funders). The second, and less often used, approach is to plan impact evaluation efforts during the planning of an educational program (Arnold, 2015). To conduct required impact evaluations, Extension educators must have and demonstrate levels of knowledge and skills about evaluation that go beyond what they have traditionally used to demonstrate participant satisfaction and gains in knowledge and skills (Kluchinski, 2014).

Evaluation Capacity Building (ECB) in Cooperative Extension

Given higher expectations for evidence to support Extension educational program accountability and quality--in the forms of behavioral and economic impacts, public value, and more—there has been increased investment in professional development for both program planning and evaluation capacity building (Arnold, 2015). Researchers have noted that an *intentional*, focused plan for Evaluation Capacity Building must be in place to support the integration of high-quality program impact assessment and scholarly reflexivity into ongoing evaluation efforts in Extension (Douglah, Boyd & Gunderson, 2003; Vengrin, 2016).

Evaluation Competencies

While much research to date has explored levels of evaluation within Extension, there is a lack of understanding and agreement regarding the specific evaluation competencies that

Extension educators must have in order to perform education program impact evaluation (Ghimire & Martin, 2015; Vengrin, 2016; Wells-Marshall, 2012). Competencies can serve as guides for both individuals and organizations to direct efforts at ECB, can help to identify gaps (Vengrin, 2016), and facilitate self-assessment (Ghimire & Martin, 2015).

Competency specification can serve as a foundation for communication about how best to plan for and do impact evaluation. The American Evaluation Association (AEA) identified a set of competencies for professional evaluators in 2018. These evaluator competencies may have implications for Extension educator/evaluators who could benefit from identifying their own evaluation competency needs to determine strengths and gaps in their knowledge. The AEA Evaluator Competency Domain areas include a) Professional practice, b) Methods, c) Context, d) Planning and Management and e) Interpersonal skills.

Evaluation Culture

The investigation of subjective norms pertaining to impact evaluation can illuminate the nuances of evaluation culture within Extension. Researchers have posited that collective social expectations regarding evaluation within the organization may influence Extension educators' desires and intentions to complete evaluations that are mandated by administration as well as for their own use (McClure et al., 2012; Morford et al., 2006a; Workman & Scheer, 2012).

Attitudes Toward Evaluation

Wells-Marshall (2012) examined factors that influence an individual's attitudes and behaviors regarding systematic program evaluation within Extension. She found that the communication of information, attitudes, and behaviors regarding program evaluation were predictors of an organizations' willingness to engage in or learn about evaluation (p. 2).

Vengrin (2016), in her study of Extension educators, found that attitude had a significant relationship to study participants' perception of the importance of specific evaluation competencies. She also found that attitudes were significant in their effect on subjective norms and on intentions to complete evaluation behaviors, which concurred with findings of previous researchers (Ajzen & Fishbein, 2005; Ajzen & Madden, 1986; Fazio, 1990).

Urban, Burgemaster, Archibald, and Byrne (2015) addressed attitudes toward evaluation in general in their study of the quality of program evaluation. They found that two factors correlated with higher or lower quality program evaluations—both optimism regarding evaluation and the valuing of evaluation.

Very little research on attitudes toward *impact* evaluation in Extension was found in the literature. Studies by Lamm (2012), Vengrin (2016), Workman (2010), and Wells-Marshall (2012) addressed attitudes toward evaluation in general within Extension but did not specifically address impact evaluation.

Impact Evaluation in Extension

Roche (1999) contended that the process of impact assessment needs to be integrated into all stages of educational programs – from planning, to implementation, to evaluation. His basic premise is that "impact assessment should not refer to the immediate outputs of a project or program, but to any lasting or significant changes that it brought about," (p. 20-21). Another researcher noted that "program evaluation, and specifically, program impact evaluation, means different things to different people including Extension administrators, content specialists, and outreach professionals," (Wells-Marshall, 2012, p. 7).

Extension program impact evaluation has been studied by researchers including Arnold (2006); Bennett (Bennett, 1975; Rockwell & Bennett, 2004); Kluchinski (2014); Kelsey (2008);

and Lamm & Israel (2013). In a review of evaluation research published in the Journal of Extension between 1965 and 2009 Workman and Sheer (2012) found that most evaluations were at the rudimentary levels of basic learner satisfaction and short-term learning. Further, Lamm and others (2013) found that, despite professional development efforts, Extension educators continue to report only "basic information on contacts made and reactions to programs, rather than on behavior changes (medium-term) and SEE [social, economic, environmental] condition (long-term) changes," (p. 4-5).

Several authors have provided insight into barriers to fully embracing and integrating assessment into Extension work, including a lack of motivation based on external demands for accountability, a perceived lack of time and/or skill to conduct assessment, and a perceived lack of support and direction regarding their role in evaluation of program impacts (Olson, Skuza, Blinn, 2007; Wells-Marshall, 2012). Vengrin (2016) found that Extension educators' attitudes about evaluation influence their intentions to conduct evaluations and, their resultant evaluation behaviors. Further, culture of the Extension organization affects Extension educators' perception of the importance of evaluation and, in turn, influences evaluation behaviors (p. 122).

Research regarding behavior has long suggested that an individual's attitudes, beliefs, and expectations are important predictors of individual behavior (Ajzen & Fishbein, 1980; Bandura, 1977b). However, one must also consider the effects of the context in which the individual is embedded regarding a specific target behavior.

Ghimire and Martin (2015) found that Extension educators reported needs for training to build evaluation capacity in these competency areas: "assessing impact of a program, developing and implementing surveys, analyzing and interpreting survey results, and using impact data for

further planning," (p.11-12). Their findings have implications for designing and implementing impact evaluation competency development programs.

Impact evaluation planning, capacity building, and implementation are complex, interdependent, multi-faceted and dynamic and may best be understood using a comprehensive, systems approach (Urban, Burgemaster, Archibald, & Byrne, 2015; Trochim, Urban, Hargraves, Hebbard, Buckley, Archibald, & Burgemaster, 2012). Comprehensive studies that address evaluation competencies, evaluation culture, evaluation capacity building efforts, and self-reported evaluation behaviors can be helpful to both administrators and leaders who are working to support the evaluation activities of their colleagues and to Extension educators at all levels of the organization who seek to strengthen their own capacities.

Lamm and Israel (2011) concluded that in-depth research to determine how some Extension professionals have been successful in evaluating their programs for impact and public value, at a higher level of rigor, and to examine how they have overcome barriers to evaluation implementation, could further enhance understanding of evaluation practice in Extension (p. 60).

Further, there is a noticeable gap in literature and calls for studies that employ theories of attitude and behavioral change to explore how various attitudes and beliefs about evaluation and assessment may ultimately impact related behaviors (Keeling, Wall, Underhile, & Dungy, 2008).

Purpose of the Study

The purpose of this study was to employ a comprehensive, systems approach to identify potential relationships among the individual and organizational factors that influence the practice of impact evaluation (IE) in Cooperative Extension. Factors included were a) perceived skill levels and importance of impact evaluation competencies; b) impact evaluation culture of the organization; c) impact evaluation behaviors and their antecedents; and d) individual and

organizational demographic factors such as educational level, program area, years of Extension experience, training, program planning models used, amount of teamwork, and more.

The methodology was primarily quantitative and employed an electronic survey of current professional staff in six states in the North Central Region of the Extension service. Results show Extension professionals' perceptions of which impact evaluation competencies are most important for them to be effective in their impact evaluation work and, also, self-assessed levels of those competencies. Further, the context of the Extension organization was examined using the Theory of Planned Behavior as a scaffold to determine factors, including organizational impact evaluation culture, which may positively or negatively affect Extension educators' impact evaluation intentions and practices.

Research Question

The research question explored in this study was: What are the relationships among the individual and organizational contextual factors that influence impact evaluation behaviors within Cooperative Extension?

The constructs and factors explored included: a) Theory of Planned Behavior factors: Attitude, Perceived Subjective Norm, Perceived Behavior control (self-efficacy, contextual constraints), behavioral intent, and target behavior (actual impact evaluation behaviors); b) Impact Evaluation competencies by perceived skill level and perceived importance; c) Impact Evaluation Culture; and d) Individual and organizational demographic factors including impact evaluation training, types of impact evaluation training, years in the organization, Extension program area, type of position (administrator, specialist, county-based), highest educational degree, support for impact evaluation, teamwork, location (state), program planning model(s), and role(s) in impact evaluation.

Significance of the Study

This study contributed to the knowledge base of evaluation competencies and related factors that may influence evaluators' ability to do quality impact evaluation. The study provided empirical evidence about evaluator competencies used specifically for impact evaluation of educational programs in Extension. Further, the study identified relationships among the many individual and organizational factors involved in performing effective impact evaluation.

Results of this study also contribute to the evaluation research conducted about context. Alkin, Christie and Vo (2012) acknowledged that many components of evaluation had been studied but a topic that needing more in-depth exploration, was that of context. King and Stevahn (2015) recommended looking at evaluator competencies used in various contexts.

Need for the Study

Evaluators and organizations that require impact evaluation can better support effective evaluation practices by increasing their understanding of important evaluation competencies, evaluation behaviors of personnel, and their organizational evaluation culture. Gaining insight into these factors can help Extension administrators more effectively convey the need for impact evaluation to professional staff. Study results may also be used to design staff development and evaluation capacity building efforts to promote higher quality, focused, and more appropriate evaluations of educational program impact and outcomes.

Definitions of Key Terminology

For the purpose of this study the following terms are defined:

Administrator (CE): Any individual that supervises more than one other Extension Educator within the Cooperative Extension service (Wells-Marshall, 2012; Vengrin, 2016).

Attitude: A disposition, based on beliefs, to respond favorably or unfavorably to a specific behavior; attitude is may predict and explain human behavior. (Ajzen & Fishbein, 1999).

Competencies: A set of skills, attitudes, and qualifications endorsed by a professional organization as guidelines and standards for practitioners (Vengrin, 2016).

Cooperative Extension: Hereafter "Extension," is a non-formal, publicly funded, worldwide educational organization that extends the research and educational resources of departments of agriculture, land-grant universities, regional, and county Extension offices to the public (Fouts, 2004; Vengrin, 2016).

Endogenous Variable: In structural equation modeling (SEM) an endogenous variable is explained by other exogenous or endogenous variables. Endogenous variables may only be dependent or may be both independent and dependent in a causal path (Kline, 2016).

Evaluation: Evaluation involves assessing the strengths and weaknesses of programs, policies, personnel, products, and organizations to improve their effectiveness (Wells-Marshall, 2012; Guba & Lincoln, 2001).

Evaluation Capacity Building (ECB): An intentional process utilized by an organization to increase individuals' skills, knowledge, and motivation to plan, conduct, and use evaluation results (Preskill & Russ-Eft, 2016; Labin et al., 2012).

Evaluation Competencies: Competencies needed to complete evaluation behaviors such as knowledge, skills, behaviors, and attitudes (Labin, 2014).

Evaluation Culture: Labin, Duffy, Meyers, Wandersman & Lesesne (2012) defined evaluation culture as the "collective values, attitudes, goals, and practices that can support or hinder organizational change as related to evaluation," (p. 5).

Exogenous Variable: In SEM, exogenous variables are independent variables, are considered antecedents or contributing causal factors, and are assumed to be determined by causes outside the core model which consists primarily of endogenous variables (Kline, 2016).

Extension Educator: Any professional of Cooperative Extension working to plan, implement, and evaluate educational programs (Vengrin, 2016).

Extension Specialist: Faculty members with expertise and specialized knowledge in a discipline or subject-matter area. Extension Specialists interpret, translate, and disseminate research-based information to county Extension agents and their clientele. Specialists generally have academic rank equivalent to their colleagues in the Land Grant campus professorial system (Seevers, Graham, & Conklin, 2007; Wells-Marshall, 2012).

Program Impact Evaluation: In this research, the terms "program impact assessment" and "program impact evaluation" are synonymous. They refer to the social, economic, and/or environmental (SEE) effects of a program. Impacts are generally long-term outcomes and may be positive, negative, or neutral; intended or non-intended (Roche, 1999).

Impact Indicator: Expression or indication of impact. This is evidence that the impact outcome goal has been achieved (Roche, 1999).

Latent Variable: Also referred to as a *factor* or *construct*, a latent variable cannot be observed. Thus, it is measured using indicators to specify the construct (Kline, 2016).

*Observed Variable**:*** Observed variables can be directly measured and often have associated measurement errors, while latent variables do not. Data for these variables is measured and exists in data files; they can be discreet or continuous variables (Glen, 2019).

Organizational Culture: A "complex set of values, beliefs, assumptions and symbols that define the way in which a firm conducts business" (Barney, 1986, p. 657).

Stakeholders: A person or organization partner, internal or external, with an active interest in the organization and its programs; an investment in the organization/program (time, money); and/or a commitment to the organization's and/or program's success (Wells-Marshall, 2012).

Subjective Norm: Social factors relating to how an individual perceives the social desirability of the behavior within the context that the behavior is expected to be exhibited (Ajzen, 1991).

Conceptual/Theoretical Framework: Factors Influencing Extension Impact Evaluation

Ravitch and Riggan (2016) asserted that "a conceptual framework is an argument about why the topic one wishes to study matters, and why the means proposed to study it are appropriate and rigorous," (p. 5). Further, they described a theoretical/conceptual framework as a superstructure for linking the elements of the research process including the literature review, methodology, and analytic approach to guide the researcher as they work to answer the research questions. Within the superstructure foundation, a researcher can meld formal and informal theory, methodology, context and setting, researcher interests and goals, and researcher identity and positionality (p. 5). Finally, they noted that a conceptual framework evolves as research progresses to accommodate learning during the process.

Extension is a complex, multi-faceted system with diverse funding sources; various government and external partners; and diverse evaluation requirements. It involves an equally complex set of evaluation capacity building (ECB) factors, relationships, and influences (Urban, Burgemaster, Archibald, & Byrne, 2015; Vengrin, 2016; Wells-Marshall, 2012). These merit study to explore relationships among the factors that contribute to doing impact evaluation successfully. Factors relevant to program impact evaluation in Extension include a) attitudes

toward impact evaluation – values, beliefs, subjective norms; b) organizational culture, norms, and support for program impact evaluation efforts; c) perceived behavior control, including self-efficacy for impact evaluation behaviors; d) impact evaluation competencies (knowledge, skills, and expertise); e) impact evaluation plans (behavioral intentions); f) actual impact evaluation behaviors exhibited by Extension educators; and g) individual and organizational demographic factors that affect the core factors (a through f) described here.

Foundational Theory and Relevant Literature

Factors included in the conceptual model are among those identified in previous research (Jones, 2015; Preskill, 2014; Vengrin, 2015; Workman & Scheer, 2012; Wells-Marshall, 2012) and in Extension assessment literature (USDA NIFA, 2017). Determining relationships among these factors may provide a better understanding of how to catalyze more effective impact evaluation behaviors through intervention via professional development, and other means.

Ajzen's Theory of Planned Behavior (TPB) serves as the core foundation for this research. The theory's fundamental concepts, shown in blue in the conceptual model (Figure 1-1) include attitude, perceived behavior control, behavioral norms, intention to perform a behavior, and the actual/target behavior. The TPB is an established theory used to explain behavior (Ajzen, 2001; Ajzen & Fishbein, 2005). Its creator has endorsed expansion and adaptation of the theory to include related factors which may contribute to determining relationships among those factors.

Organizational culture and organizational learning are related concepts which affect the successful implementation of change and innovation within an organization (Schein, 1992; Cameron & Quinn, 1999). Schein (1992) defined group culture as a "pattern of shared basic assumptions that the group learned as it solved its problems of external adaptation and internal integration, that has worked well enough to be considered valid and, therefore, to be taught to

new members as the correct way to perceive, think, and feel in relation to those problems" (p. 12). Berrio (1999) noted that organizational learning has potential to influence behavior based on knowledge and insights gained. These concepts and related theories have implications for the successful integration of effective impact evaluation behaviors into Extension. Further, Berrio found that Extension exhibited a culture that valued teamwork, had a high level of commitment to the organization, and valued both individual and organization development.

Role theory literature presents conflicting research on whether filling multiple roles results in positive or negative effects on individuals' psychological well-being (Biddle, 213). Van der Horst (2016) noted that, overall, role theory research regarding the effects of filling multiple roles has shown that positive results are more likely than negative. Marks (1977) posited that level of commitment affects and is affected by multiple role participation.

Preliminary Theoretical Model

The concept map in Figure 1-1 illustrates the factors included in this research that shape the context in which impact evaluation in the Cooperative Extension Service must be performed.

This concept map was created based on theories and related concepts found in the literature including the Theory of Planned Behavior, capacity building theory, learning organization theory, role theory, evaluation theory, program planning models, program evaluation and impact evaluation models, and related research. The factors included were judged to have possible influence on professionals' ability to complete educatonal program impact evaluation within the Cooperative Extension Service context. The researcher's own experience planning and performing both traditional and impact evaluation of educational programs in Extension was also a source of input for the prelimnary theoretical model.

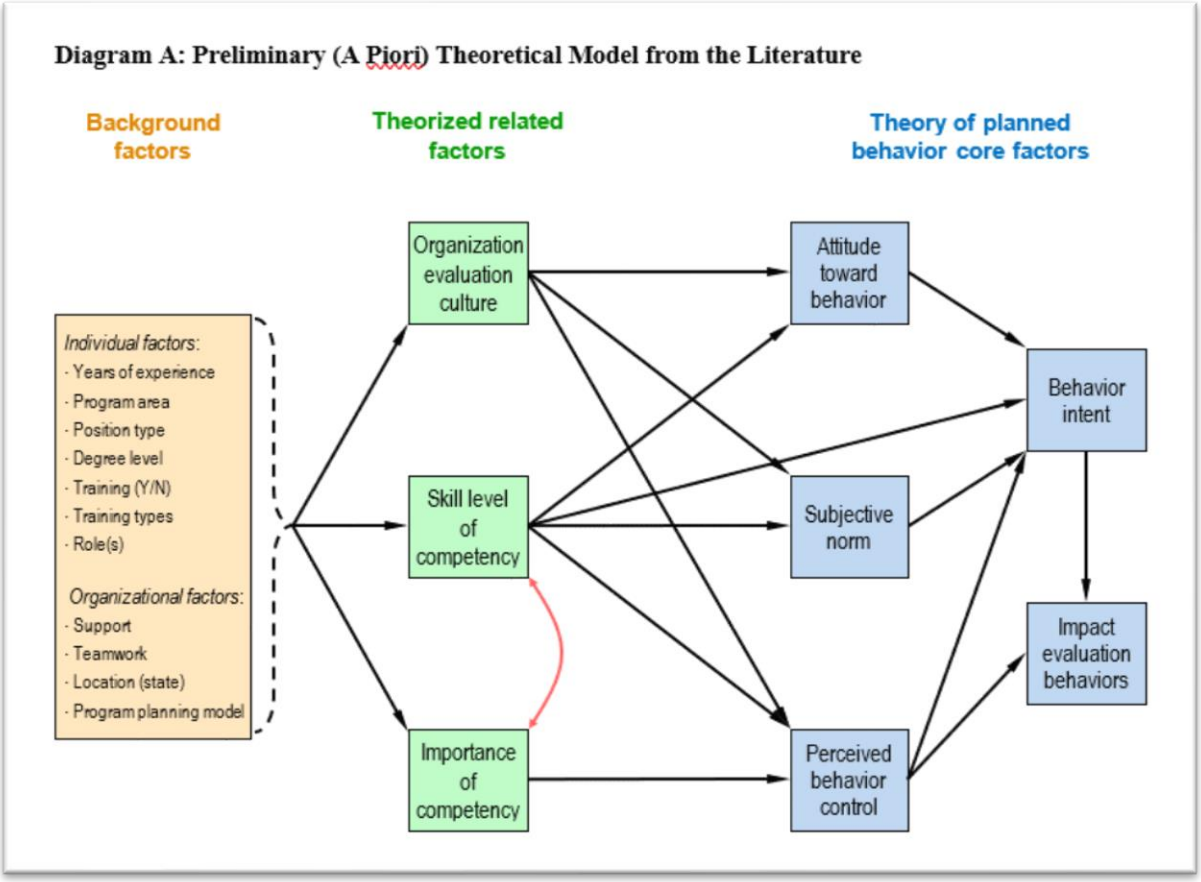

Figure 1-1. Theoretical model developed from the literature. Black arrows represent direct effects (pathways). The red double-headed arrow indicates correlated disturbance (error) terms.

The concept model shows the relationships and potential relationships among the factors and illustrates the interdependence of the factors within the complex Extension system. The model was created based on the theories and literature outlined previously, as well as on experiences of the researcher. Determining the strengths of these relationships was the focus of the research. The research question addressed the relationships among the factors and variables included in the theoretical concept model and guided specification of an initial model which was the basis for data analysis using structural equation modeling (SEM).

The fundamental concepts in the theoretical model include attitude, perceived behavior control, behavioral norms, intention to perform a behavior, and target behavior which are characteristic of the Theory of Planned Behavior (TPB), an established theory used to explain

behavior (Ajzen, 2001; Ajzen & Fishbein, 2005). Additional factors chosen based on review of academic literature include a) organizational impact evaluation culture; b) impact evaluation competencies by perceived level of skill and perceived importance; c) individual demographic background factors including impact evaluation roles, years of Extension experience, highest academic degree, position type, program area, training in impact evaluation, and types of training; and d) organizational demographic factors including teamwork, program planning model(s) used, support for evaluation, and location (state).

Examining these factors provides a comprehensive view of impact evaluation behaviors and impact evaluation culture within the Extension organization. Vengrin (2016) noted that most literature relevant to evaluation behaviors in Extension does not consider a specific set of evaluation competencies. Evaluation culture is considered in the literature but not specifically for impact evaluation. A comprehensive approach which considers competencies, evaluation culture, and other potentially influential factors is possible by expanding the Theory of Planned Behavior and using SEM to analyze the effects among the factors.

Data collection employed previously validated items from Vengrin's (2016) Evaluation Competencies, Culture and Behavior (ECCB) instrument; items from Preskill and Torres' ROLE subscale on organizational culture (2001); items adapted from guidelines on how to measure constructs in the Theory of Planned Behavior model (Ajzen & Fishbein, 2005); and demographic items. Data analysis produced descriptive statistics (mean, medium, mode, etc.) for demographic and Likert-scaled items; composite scores for competency, behavioral, and culture items; structural equation modeling (SEM) path analysis of factors included in a revised conceptual model; thematic analysis of qualitative items; and multiple regression (ANOVA) to analyze factors not included in the re-specified final SEM model.

Summary: Research Need, Goals, and Potential Implications

The demand for evidence of behavioral impacts of Extension educational programs has increased over the past two decades. Extension administrators cannot assume that Extension state specialists, regional and county agents agree with their organizational goals regarding impact evaluation capacity building (Wells-Marshal, 2012; Minarovic & Mueller, 2000). To increase support for evaluation capacity building and for the performance of desired impact evaluation behaviors, it is important to measure and understand the individual mindsets and competencies of those within the organization.

Behavioral research has shown that an individual's attitudes, beliefs, and expectations are important predictors of individual behavior (Ajzen & Fishbein, 1980, 1991; Bandura, 1977b). Demographic factors such as experience, staff development and training, academic background, and other factors may also influence both individuals' intent to perform a behavior and their actual performance of that target behavior. One must also consider the effects of the context in which the individual is embedded regarding a specific target behavior such as impact evaluation. Context includes organizational culture; includes internal and external influences on individuals; and perceptions of organizational support for impact evaluation.

A preliminary conceptual theoretical framework based on the Theory of Planned Behavior, capacity/ competency building literature, learning organization theory, role theory, Extension literature and the researchers' experience was created. The a priori theoretical concept model was specified to guide the research using structural equation modeling (SEM) as the methodology. This design served as the foundation for the literature review, methodology (instrumentation, data collection, and analysis), and interpretation of the results.

Research goals included determining and understanding relationships between the many factors that influence program impact evaluation behaviors in Extension. Potential implications of the research were to a) identify important impact evaluation competencies; b) determine targets for ECB efforts; c) identify factors that affect actual evaluation behaviors and practice; and d) identify desirable Extension impact evaluator dispositional factors.

Research results are presented in chapter four and discussed with respect to contributions to theory, implications for practice, and recommendations for future research in chapter five.

CHAPTER 2. LITERATURE REVIEW

The purpose of this study was to employ a comprehensive, systems approach to identify potential relationships among the individual and organizational factors that influence the practice of impact evaluation (IE) in Cooperative Extension. Factors included were a) perceived skill levels and importance of impact evaluation competencies; b) impact evaluation culture of the organization; c) impact evaluation behaviors and their antecedents; and d) individual and organizational demographic factors such as educational level, program area, years of Extension experience, training, program planning models used, amount of teamwork, and more. The literature review provides background on the factors and interrelationships of those factors found in literature. The preliminary theoretical and conceptual framework for the study was introduced in chapter one and serves as the foundation for the literature reviewed in this chapter.

Influences on Impact Evaluation Behaviors in Cooperative Extension

This literature review provides an overview of the history and current status of evaluation in Cooperative Extension (Extension); summarizes factors that have contributed to the current focus and increased pressure for accountability, impact and public value documentation; introduces the construct of evaluation competencies; outlines evaluation capacity building efforts in Extension; reviews research on internal evaluator roles and related issues; and presents factors influencing the success of Extension educators in integrating and performing high-quality program impact assessment.

The underlying bases for the preliminary theoretical framework for examining current impact evaluation competencies of Extension educators and the factors that influence impact evaluation practice in Extension, including organizational evaluation culture, impact evaluation-related attitudes, demographic factors, and impact evaluation behaviors are included.

Brief Overview of Program Evaluation

Evaluation is an integral component of many professions, particularly those whose work is funded through various public and private funding agencies (Roche, 1999; Volkov, 2011). Grantors, private and government funding entities increasingly require evidence--reported in the forms of program outcomes, impacts, and public value--in order to determine how well funding is being utilized (Kalambokidis, 2011). In an era of rapid change, complexity, and accountability, evaluation is critical to both program and organizational effectiveness, as well their survival (Argryis & Schön, 1978; Preskill & Russ-Eft, 2016).

For evaluation work to be effective, individuals need to have a clear understanding and positive attitude toward the field and practice of evaluation (Lekies & Bennett, 2011; Parkinson, 2009; Radhakrishna & Martin, 1999). Individuals often view evaluation negatively and, as a result, are resistant to undergoing or implementing it. Evaluation competencies--knowledge, skills, and behaviors—are also essential to effective evaluation of educational program impacts.

History of Evaluation Approaches in Cooperative Extension

Nichols, Blake, Cazdon, and Radhakrishna (2015) provided an historical overview of evaluation in Extension, highlighting what they judged to be the focus on evaluation efforts during periods they identified. Their analysis has much to offer. Below is a summary of the eras of evaluation they identified for the past century, key developments, and the priorities and goals of Extension evaluation efforts during each period.

1914 to 1976 – Assessing Operations, Activities, and Participation

The Smith-Lever Act of 1914 established the Extension system (Association of Public and Land-Grant Universities [APLU], 2012). Per the Smith-Lever Act, the U.S. Congress requested “a full and detailed report of its operations” from Extension. The impact was not

reported nor an explicit concern. Reporting was not mandatory during this period; however, some advocated for it:

> A major challenge confronting the educational institution is that of determining the impact of its planned programming efforts in effecting desired behavioral changes in its publics. To achieve this end every subsystem within the organization must perfect and utilize tested and valid methods for pinpointing evidences in relation to their program objectives and for collecting such evidences. (Boone, Dolan & Shearon, 1971, p. 18)

1977 to 1989 – Focus on Accountability

Rennekamp and Engle (2008) addressed new attitudes that brought about a new era for Extension program evaluation by including the quote below in their article (p. 22):

> No longer can it be taken for granted that programs are good and appropriate. Extension is operating in a new environment—an environment of more open criticism and demands for justification of actions. All publicly funded agencies, not just Extension, are vulnerable in these times. In an era of accountability, Extension must be able to defend who and how people are being served. It also needs to document that programs are achieving positive results (Andrews, 1983, p. 8).

During this time period, Extension organizations began to hire evaluation specialists, according to Lambur (2008). These individuals often focused on evaluation capacity building such as training and assisting staff, rather than functioning as dedicated evaluation specialists who conducted evaluations (Guion, Boyd, & Rennekamp, 2007; Rennekamp & Arnold, 2009).

1990 to 2005 – Shift to Evaluating Measurable Outcomes

The transition from accountability reporting to outcomes measurement began at the state level in the early 1990s. Around the same time, evaluation on the federal level changed when a

national task force on accountability and evaluation called for system-wide outcome accomplishment data. This, along with decreases in federal, state, and local funding, prompted the passing of the Government Performance and Results Act of 1993 (GPRA), which:

> … requires each federal program to identify indicators of outcome for major programs, to provide targets at the beginning of each fiscal year for each indicator, and to report on the actual values for each outcome indicator within six months after the end of the fiscal year, (Hatry, 1997, p. 32).

The U.S. government enacted the GPRA to "focus government activity on results, rather than on inputs or process" (Wargo, 1994, p. 65).

2006 to 2015 – Multistate, Regional, and National Evaluation Collaboration

In a review of evaluation studies that were published in the *Journal of Extension* from 1998 to 2007, Duttweiler (2008) observed evidence of increased evaluation collaborations and teamwork at state, multicounty or instate regional, and multistate levels.

Across all areas of programming, Nichols, et al. (2015) found a consistent theme of Extension building strong relationships with individuals in organizations and their communities and postulated that the relationships have, indeed, influenced important changes. Thus, program impact evidence showed that social capital was facilitated and that "features of social organization, such as networks, norms, and trusts that facilitate coordination and cooperation for mutual benefit" by Putnam (1993, p. 35). Nichols, et al. (2015) contended that Extension had not done an adequate job of documenting those impacts and needed to increase the capacity to do so. To document program worth, Extension educators must first improve their effectiveness at providing evidence of its private value to those directly served by its programs (Nichols, et al., 2015, p.98). Further, in order to embrace the emerging focus of how Extension programming

creates public value, a type of impact that goes beyond serving traditional clientele, Extension must address the value to those who do not participate in its programs (Kalambokidis, 2004).

Current Focus of Evaluation in Extension

The current focus of evaluation in Extension is on providing evidence of the impact of programs to those who directly participate (private value) and considering what value those program impacts have to the public (public value) (Kalambokidis, 2017). Rennekamp and Engle (2008) concluded that an ongoing result of outcomes/impact evaluation has been to "entrench behavioral change as a logical and valued outcome of Extension programming," (p. 20).

Ghimire and Martin (2015) observed that Extension educators' assessment of educational program success is often limited to short-term outcomes. This is accomplished by measuring immediate reactions to a program and measuring increases in knowledge and skills using pre- and post-tests (Franz & Townson 2008). Assessment of program impacts in the form of actual behavior changes over time are less frequent; these include measuring social, economic, and environmental (SEE) impacts (Lamm, Israel, and Harder 2011). Ghimire and Martin (2015) suggested that evaluation capacity building (ECB) be demand-driven and address competencies judged as important to Extension educators' efforts to engage in program impact evaluation.

Impact Evaluation and Assessment

Impact assessment and evaluation are, essentially, about "the measurement and valuation of change" (Roche, 1999, p. 24). Such evaluation is generally aimed at one or a combination of interventions designed to affect change in the lives of people, communities and societies. Impact assessment has been performed for decades to show results, for accountability purposes, for program improvement, and for future planning (Kellogg Foundation, 2004; United Way of America, 1996; World Bank, 1997).

Roche (1999) shared this definition of impact assessment, which was based on the work of a joint action research project completed in the 1990s by several international, non-government organizations (NGOs), including Oxfam of Great Britain, Novib of the Netherlands, and World Bank collaborators:

> Impact assessment is the systematic analysis of the lasting significant changes – positive or negative, intended or not – in people's lives brought about by a given action or series of actions (p. 21).

Roche argued that the process of impact assessment must be integrated into all stages of educational programs – from planning to evaluation (p. 18). A systems approach, which includes clarification of impact outcome goals, identification of indicators of change to be measured throughout a project, methods for measuring and collecting indicator data, and plans for management and reporting of impact evidence are all part of the process (Patton, 1982; Roche, 1999; United Way of America, 1995).

Impact evaluation is a complex undertaking. It requires careful planning to ensure that impact evidence indicators are identified early in the program planning process and are monitored in a planned, systematic manner (Roche, 1999). The Kellogg Foundation (2004) offers the following points of consideration when deciding on an evaluation design: "create a flexible and responsive design; collect and analyze information from multiple perspectives; and always return to your evaluation questions" (p. 70).

Program impact evaluation requires data collection; however, "methods follow purpose" (Taylor-Powell & Steele, 1996, p. 1). Prior to an impact evaluation process, it is critical that these types of concerns be addressed: a) the purpose and the rationale of the evaluation; b) who is the target audience and how will they use the information; c) what information do stakeholders want and/or need to know; d) what are the essential questions the impact evaluation must answer; and e) what are the best sources of evidence (information) (Mohr,1995; Patton, 2003;

Powell & Steele, 1996; Russ-Eft & Preskill, (2016); Taylor-Powell & Steele, 1996; USAID Center for Development Information and Evaluation, 2017; Wells-Marshall, 2012).

Patton (2008) reported that "appropriateness" is the standard affirmed by the American Evaluation Association position statement on "scientifically based evaluation methods" (AEA, 2003). The European Evaluation Society also stressed "the importance of a methodologically diverse approach to impact evaluation" (EES, 2007). Further, the Network of Networks on Impact Evaluation (NONIE) advocated a standard of appropriateness concerning impact evaluation. NONIE was established by international evaluation offices representing the United Nations, World Bank, other development organizations, and developing country representatives to provide guidance for conducting impact evaluations as outlined in this key statement:

> Methods, techniques and approaches for impact evaluation should match the specific circumstances of the evaluation—its purpose, the nature of the intervention, the questions, the level of existing knowledge, and the resources available. *Methodological appropriateness should be considered the "gold standard "for impact evaluation (*NONIE, 2007).

Program impact evaluation has evolved since it began in the 1950s to go beyond assessing observable program results to involving local stakeholders to obtain input through a more participatory process (Roche, 1999). Wholey (1987) and Patton (1989) emphasized the role of stakeholders and program staff in a utilization-focused approach to program planning and evaluation. Core tenets of the process focus on the measuring of change in one or more of the following areas: environmental impact assessment (EIA), cost-benefit analysis (CBA), social impact assessment (SIA), and social-cost-benefit analysis (SCBA), among others (Howes, 1992; Leeuw & Vaesen, 2009).

Program Impact Evaluation in Cooperative Extension

Extension professionals are being asked to provide stronger evidence of educational program impacts. They must collect data, analyze and report it to show that programs contribute

to learning which, in turn, causes positive impacts in clients' personal and professional lives, and in their communities. Focus on increased accountability is a result of pressure to demonstrate relevance in an era of declining funding (Decker & Yerka, 1990; Kelsey, 2008; O'Neill, 1998; Rennekemp & Engle, 2008). Impact documentation needs stem from Federal compliance requirements and from a need for publicly available outcomes data for institutional and state comparisons (Kelsey, 2008). Extension must demonstrate a positive return-on-investment to clientele and the public (Franz, Arnold, & Baughman, 2014). Program impact evaluation must not only be an accountability-driven add-on but a fully integrated, essential element of practice. Assessment of learning and the impacts of that learning are fundamental to this work and, in the words of Astin & Associates "assessment is not an end in itself but a vehicle for educational improvement," (Astin, Banta, Cross, El-Khawas, Ewell, Hutchings, & Moran, 1992, p. 1).

Program impact evaluation is important for accountability, decision making, and program sustainability in Cooperative Extension (Marshall, Higginbotham, Harris, & Lee, 2007). A 1980 report mandated by the Food and Agriculture Act of 1977 found accountability work by Cooperative Extension professionals to be "short on impacts," (Warner & Christenson, 1984, p.17). Much progress has been made; however, a review of evaluations reported in the Journal of Extension found that many were at a level no higher than learner satisfaction and short-term learning (Workman & Scheer, 2012). Lamm, Israel and Diehl (2013) noted that, despite professional development efforts, Extension educators continue to report only "basic information on contacts made and reactions to programs, rather than on behavior changes (medium-term) and SEE [social, economic, environmental] condition (long-term) changes," (p. 4-5).

Wells-Marshall (2012) found that Extension state and field staff are committed to using evaluation results, to data analysis, and to focusing evaluations. Communication of information, attitudes, and behaviors regarding program evaluation were found to be predictors of an Extension organizations' willingness to engage in or learn about evaluation (Wells-Marshall, 2012). This willingness may support capacity building for impact evaluation.

Patton (2008) noted that context is important in Extension educators' abilities to complete effective program impact evaluation. Context involves many variables, such as individuals' perception of the value of the task of impact evaluation, individuals' perceived level of evaluation competencies needed to do impact evaluation, perceived organizational climate, perceived evaluation culture of the organization, time available for an evaluation project, and more. It is important for evaluators to understand the context in which they perform their evaluation to ensure its results will be used most effectively (Rog, 2012). Weiss (1998) shared reasons why it is beneficial for evaluators to understand context. She indicated that evaluators need to develop a clear understanding of the factors being evaluated; to formulate inclusive and relevant questions; to assist in interpreting the data and evidence; to make informed recommendations; and to provide reporting in a useful format (p. 47).

Program Planning for Impact Evaluation in Cooperative Extension

The Kirkpatrick Model (1959, 1996, 2010), Bennett's Hierarchy (Bennett, 1967; Bennett, 1975; Bennett, & Rockwell, 1995), and the logic model (UWEX, 1995; United Way of America, 1996; Taylor-Powell & Henert, 2008) are three tools employed to give direction to Extension professionals as they plan, conduct, and report educational program impact evaluations. The emerging Extension national Impact Collaborative model further supports program design which identifies evidence needed to document various levels of program impact (eXtension, 2017).

Program planning in Extension essentially aims at documenting the following to plan and evaluate educational programs: a) determine conditions, issues, or needs; b) identify target audience and its characteristics; c) determine short-, medium-, and long-term goals and desired outcomes; e) determine content and major messages; f) determine best educational activities and strategies (outputs) to accomplish educational goal(s); g) outline resources needed (time, materials, and funding); h) establish a program promotion and marketing plan; i) create a timeline for

implementing the program; j) determine how outcomes and impact will be assessed; and k) make a plan for sharing outcomes, impact results, and successes. This planning format is based on the Logic Model originally created by professionals from the University of Wisconsin Extension system and guidelines by Taylor-Powell, Steele, and Douglah (1996) for using it. The program planning process clearly integrates evaluation planning into the process.

Diem (2003) defined impact in Extension programming as "the positive difference we make in people's lives as a result of the programs we conduct," (p. 1). To make an impact, the results of an Extension program must ultimately change people's attitudes or behavior, or benefit society in other ways (Diem, 1997). Rockwell (2002) noted that impacts are a form of program outcomes and must be planned for during program development by identifying indicators of that impact such as targeted behavior change early in the process. Changes traditionally evaluated in Extension program planning were based on what was termed "KASA" change: K (knowledge), A (attitude), S (skill), and A (aspirations) (Bennett, 1975; Rockwell & Bennett, 2004).

Evaluation Purposes in Extension

Duttweiler (2008) reviewed nine exemplary published evaluation studies of Extension educational programs and concluded that:

> … evaluation studies had influenced Extension practice by helping to establish program direction, improving existing educational practice, informing public policy, establishing or sustaining program support, offering a basis for resource allocation decisions, influencing relationships with stakeholders, and strengthening evaluation practice itself, (p. 99).

Patton (2017) recently reiterated three main purposes of evaluation: a) for *making judgments* – commonly called summative evaluations; b) for *improving programs* – commonly called formative evaluations: and c) for *ongoing development* – sometimes called developmental

evaluations. He added others, including knowledge building, meta-evaluation, lessons learned, effective practices, accountability, and monitoring. Patton further advocated that organizations infuse evaluative thinking as a type of evaluation "process use" and employ capacity-building to help focus process use. Further, he stressed to evaluators that a primary challenge in effective impact evaluation is that of matching the evaluation design to the evaluation's purpose, to the resources available, and to a realistic timeline in order to facilitate results use (Patton, 2012).

Evaluations in Extension must focus on the degree to which participants have adopted what was taught to make changes in their lives rather than evaluating the quality of Extension educators' teaching. Logic models are not only tools for program planning but, if used properly, demand that Extension set priorities. The prioritizing of outcomes gives focus to the organization, guides education programming, and requires that results be measured.

In the 1980s outcomes evaluation and the transparency that accompanies public accountability increased the stakes for those in Extension (Patton, 2008). Duttweiler (2008), Patton (2008), Rennekamp and Engle (2008), and others urged both the discipline of evaluation and the Extension service to adopt systems perspectives in their theories of change.

Many Extension-focused researchers have noted that evaluation capacity building is necessary for Extension staff because of the many competing purposes of evaluation and evaluation research. Purposes identified included program improvement, accountability, summative evaluation, monitoring, and knowledge generation (Ghimire & Martin, 2011; Powell & Boyd, 2008; Patton, 2008; Duttweiler, 2008).

Duttweiler's 2008 review of the uses of Extension evaluations published in the *Journal of Extension* from 1998 to 2007 showed that 40% reported program improvement as a primary use and 35% reported a more accountability-oriented focus of results use.

Measuring significant program impact is realized through documenting evidence of either behavior/practice change (medium-term outcomes) and/or end results (long-term outcomes)

(Diem, 2003). Impact level evaluation must go beyond assessing the traditional KASA types of change, which are considered short-term outcomes (Taylor-Powell & Henert, 2008). Extension professionals at all levels of the organization are being called upon to support their work using high-quality data to show evidence of educational program impact (Guion, 2007).

A Closer Look at Three Extension Program Planning Models

Many program and evaluation planning approaches and models are utilized to evaluate Extension programs; however, Kirkpatrick's model (1959), Bennett's Hierarchy (1975), and the logic model (1995) are the most often used. Graphic representations of each can be used to document and illustrate evidence of impact. In 1975, Bennett's Hierarchy was published for the first time in the *Journal of Extension* specifically for use by Cooperative Extension educators. Bennett's Hierarchy was created by incorporating the elements of Donald Kirkpatrick's four-levels (introduced and published during 1959) for evaluating training programs (Rockwell & Bennett, 2004). Another modern or more recent tool that has evolved from both Kirkpatrick's model and Bennett's Hierarchy is the logic model (Hoffman & Grabowski, 2004).

Kirkpatrick's Four-Level Model

Kirkpatrick's model has long served as a standard for evaluating training programs (Rajeev, Madan, & Jayarajan, 2009; Bates, 2004; Holton, 1996; Workman, 2010). The model was originally developed and published by Donald Kirkpatrick in 1959 and is well suited to provide the type of information desired by business managers (Stup, 2003). It has also been used in Extension as both an evaluation model and a foundation for newer models (Bennett, 1975; Hoffman & Grabowski, 2004). The Kirkpatrick Model was the first model developed to provide more tangible measures of impact instead of measuring only reactions or feelings of participants; it was a forerunner for Bennett's Hierarchy and the modern-day logic models.

The Kirkpatrick Model as shown in Figure 2-1 consists of four levels including: (1) reactions, (2) learning, (3) behavior (or transfer), and (4) results (Kirkpatrick, 1959, 1996, 2010).

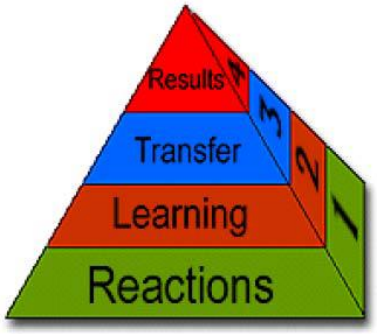

Figure 2-1. Kirkpatrick's four-level model (Winfrey, 1999).

Reaction is a measure of how participants feel about a program. Learning is a measure of the knowledge gained, skills improved, or attitudes changed due to a program. Behavior is the extent to which participants changed their behavior as a result of a program. Behavior is also referred to as transfer or application (Kirkpatrick, 1996; Winfrey, 1999). Results, the final level, are changes that happened due to a program such as increased quality, productivity, and/or profit, Kirkpatrick (1959, 1996) noted that evaluation becomes more difficult, complicated, and expensive as the levels increase; however, evaluations become more meaningful.

Bennett's Hierarchy

Claude Bennett developed his original program development and evaluation hierarchy or seven-link "chain of events" for Cooperative Extension in 1975. Levels of the hierarchy as shown in Figure 2-2 include: inputs; activities; people involvement; reactions; *k*nowledge, *a*ttitudes, *s*kills, and *a*spirations (KASA) change; practice change; and end results.

Figure 2-2. Hierarchy of evidence for program evaluation (based on Bennett, 1975).

In recent years, Rockwell and Bennett (2004) collaborated to modify and expand the original hierarchy. They added program development and modified the evaluation section to create a system called Targeting Outcomes of Programs (TOP). The purposes of TOP are to focus impact on social, economic, and environmental (SEE) issues and to document program progress (Rockwell & Bennett, 2004). A representation of the TOP model is shown below:

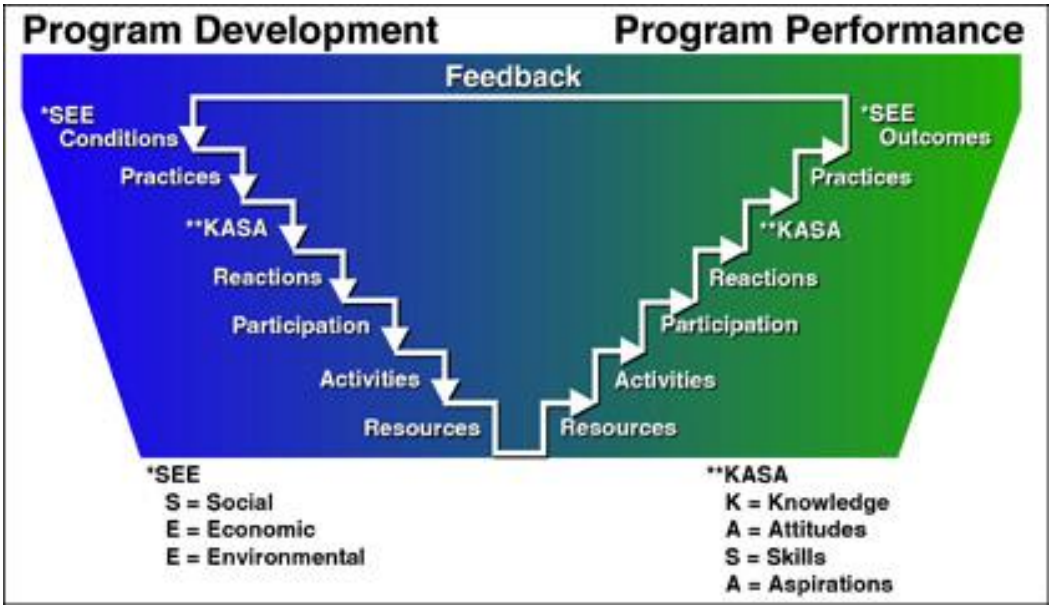

Figure 2-3. TOP (Targeting Outcomes of Programs) model (Rockwell & Bennett, 2004).

The TOP model is more useful for planning programs to account for measuring program impacts in the form of social, economic, and environmental (SEE) changes.

The Logic Model

The roots of the model are found in both Kirkpatrick's model and Bennett's Hierarchy (Hoffman & Grabowski, 2004). There are many descriptions and definitions associated with the logic model. There are also a number of other terms used to describe it, including: outcome map, program logic, idea map, action map, mental model, program action, conceptual map, and model of change (Knowlton & Phillips, 2009; Taylor-Powell & Henert, 2008; Workman, 2010).

A logic model shows the resources and the actions needed to reach a program's intended results (W. G. Kellogg Foundation, 2004). It may also be viewed as a series of relationships that represent the progression from resources to impact (Taylor-Powell & Henert, 2008).

A Simple Logic Model. In its simplest form, a logic model looks like this:

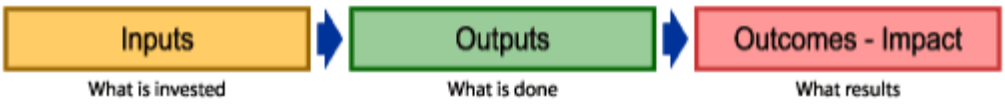

Figure 2-4. Simple logic model (UW-Extension, Madison, WI).

The graphic representation above shows the logical relationships between program resources (inputs), the educational activities that make up program delivery (outputs), and the changes or benefits that result (outcomes) (Taylor-Powell, Jones, & Henert, 2003; Workman, 2010).

The logic model describes the sequence of events designed to bring about benefits and/or change over time. It shows a logical chain of reasoning that links inputs/investments to outcomes/results. A logic model is, simply, a systems model that illustrates the connection of interdependent parts that make up the whole process and components. The current logic model being used in many Extension organizations is shown in Figure 2-5.

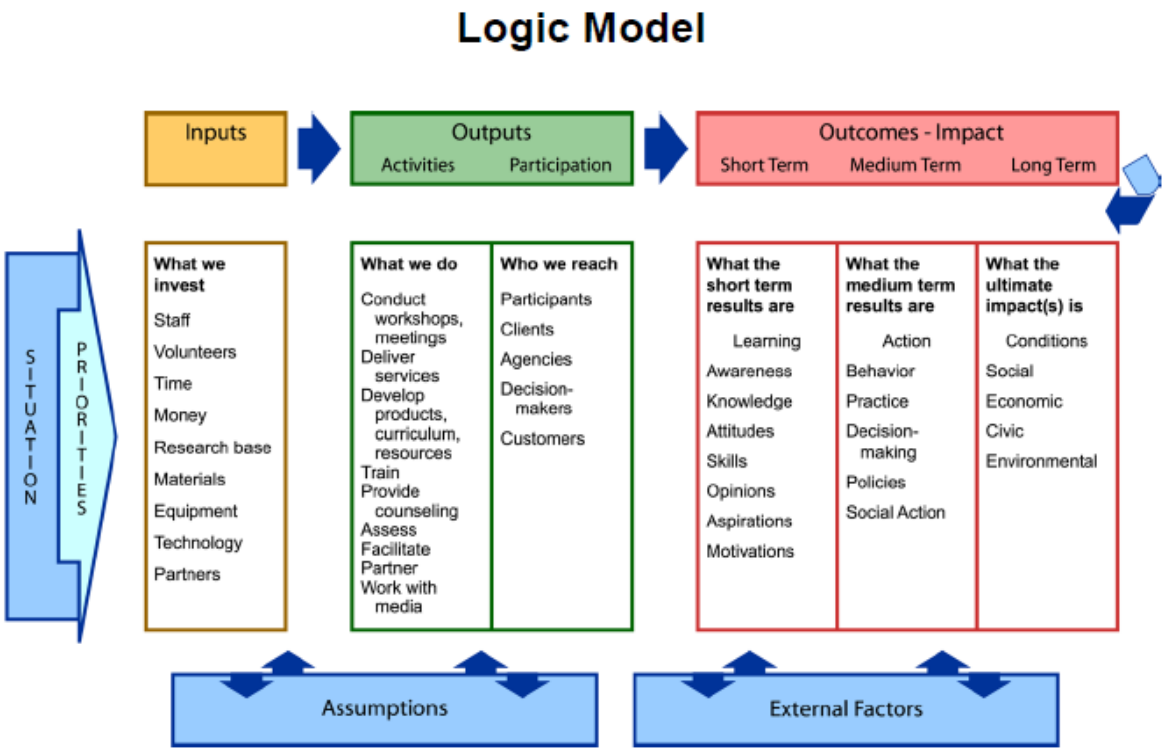

Figure 2-5. Enhancing program performance with logic models. University of Wisconsin-Extension, Feb. 2003.

The University of Wisconsin-Extension logic model shown in Figure 2-5 includes the following program components: situation/priorities, inputs, outputs-activities, outputs-participation, short-term outcomes, medium-term outcomes, long-term outcomes, assumptions, and external factors.

A logic model helps Extension educational programmers and evaluators focus on appropriate process and outcome measures. Impact in this model refers to the ultimate consequence or effects of the program—for example, increased economic security, reduced rates of teen pregnancy, and improved air quality. Impact is synonymous with long-term outcomes or goals and may include observable behavioral or social longer-term changes, as well as economic, environmental, or civic condition changes. Taylor (2007) noted that attitude and perspective transformation are some of the most difficult concepts to assess (p. 180).

In common usage impact and outcomes are often used interchangeably. Figure 2-6 shows the basic framework for a logic model and illustrates the distinction between program outputs and outcomes/impact.

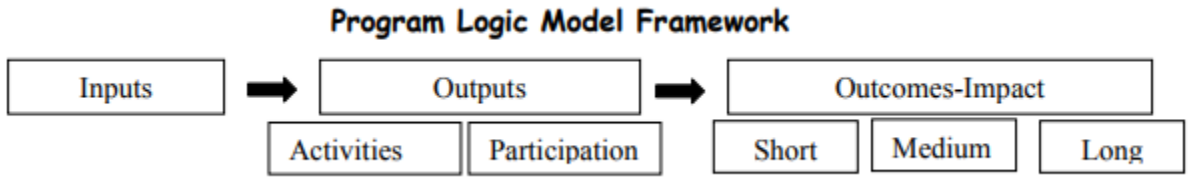

Figure 2-6. Program logic model framework (UW-Extension, Madison, WI).

A logic model is a simplified picture of a program, initiative, or intervention that is a response to a given situation. It shows logical relationships among the resources invested, the learning activities that take place, and the benefits or changes that result. Some refer to it as *program theory* (Weiss, 1998) or the program's *theory of action* (Patton, 1997). It is a "plausible, sensible model of how a program is supposed to work," (Bickman, 1987, p. 5).

Chain of Outcomes. Focusing on providing evidence of Extension program impacts requires understanding that impacts are a type of outcome and that, in program impact evaluation models, there is a "chain of outcomes" (Taylor-Powell, Jones, & Henert, 2003; Patton, 2008; United Way of America, 1996). The chain of outcomes is sequential and illustrates that outcomes fall along a continuum, from short-term, to medium-term, to long-term. The continuum is also referred to as an "outcome line" (Mohr, 1995), an "outcome sequence chart" (Hatry, 1999), or an "outcome hierarchy" (Funnell, 2000). This model concept--a series of sequential outcomes that are connected--is fundamental to a program logic model and, in essence, its simplest form.

Viewing the set of outcomes from a theory of change focus, the set of connected impacts/ outcomes may be viewed as a "pathway of change," (Earl, Carden & Smutylo, 2001). Defining outcomes as "changes in behavior" focuses program goals on measurable impacts and stresses that, to be effective, educational programs must go beyond information dissemination; they must

actively engage participants in the adoption, adaptation, and application of what is taught (p. 2). Outcome Mapping focuses planning, monitoring, and evaluation on targeted behaviors, actions, and relationships in a program's sphere of influence. This allows program developers, stakeholders, and clientele to measure impact and report evidence that supports claims of public value (Kalambokidis, 2011).

Indicators of Change. Behavior change indicators must be identified in order to obtain valid measures of program impacts. This becomes the evidence that shows stakeholders changes that can be attributed to an educational program intervention. Roche (2011) described three categories of indicators of behavior change used to show evidence of impact: material wealth, social well-being or human capital measures, and empowerment or political measures (p. 45). The type of indicator must match the impact outcome goals of the program and be measurable.

Impact evaluation, thus, is clearly integrated into the program planning models and processes employed by Extension educators.

Evaluation Capacity Building

Evaluation capacity building (ECB) is viewed as an emerging subfield of evaluation (Shaw, Shaw, I. Greene, & Mark, 2006). Preskill and Boyle (2008) noted that ECB involves strategies designed to help individuals learn "from and about evaluation," (p. 443). Labin, (2014) reported evaluation capacity building outcomes at the *individual* level included attitudes, knowledge, and behaviors and at the *organizational* level included evaluation practices, leadership, evaluation culture, mainstreaming, and resources. D'Ostie-Racine, Dagenais, and Ridde (2016) found that evaluation activities resulted in participants learning about their programs, evaluation processes, and research methodology. Concern with building evaluation

capacity has evolved to not only help educational programmers increase their *capacity to do* evaluation but also the *capacity to use* it (Cousins, et al., 2014).

For quality evaluation practice to be sustained, despite employee turnover and other possible barriers, Preskill and Boyle (2008) posited that:

> ECB needs to be supported at a cultural level (as communicated by leaders, by the organization's evaluation vision, by a living strategic plan for evaluation, and in the ways members talk about evaluation), within the performance appraisal and professional development systems, and in the organization's systems and structures (how work gets done and by whom), (p. 456).

Evaluation capacity building is seen as one component of a learning organization (Gagnon, et al., 2018). The concept of nonprofit educational organizations such as schools, universities, and government entities as learning organizations has been influenced by the ideas of Senge (1990) and Fullan (1993). Goals for transforming nonprofit organizations involve addressing the degree to which the organization has a learning culture. Individuals' and the organization's readiness for engaging in outcomes measurement evaluation and research may affect the level or resistance, negativity, and fear that can hinder effective impact evaluation (Botcheva, White, & Huffman, 2002, p. 431).

Evaluation Capacity Building in Extension

Research shows that an *intentional*, focused plan for ECB must be in place to support integration of high-quality program impact assessment into ongoing evaluation efforts in Extension (Douglah, Boyd & Gundermann, 2003; Vengrin, 2016).

Valuing the evaluator, not just the doing and reporting of evaluation, is characteristic of Extension administrators who purposefully initiate well-designed ECB strategies including mentoring, professional development, and experiential learning (Vengrin, 2016; Wells-Marshall,

2012; Fouts, 2004). Over time, providing Extension educators with opportunities for guided experiential learning may develop evaluation capacity across the organization, thus providing opportunities for organizational, as well as, individual learning (Arnold, 2006).

Extension is a complex, multi-faceted system with diverse funding sources; various government and external partners, and diverse evaluation requirements; thus, it presents a very complex set of ECB-related factors, relationships, and influences (Vengrin, 2016; Urban, et al., 2015; Wells-Marshall, 2012).

Evaluation Culture in Extension

An integral component of an organization's capacity for doing evaluation is driven by the evaluation culture of that organization (Preskill & Boyle, 2008; Labin, 2014). Arnold (2006) noted that the Extension organization has increasingly supported development of Extension educator's evaluation skills; others stressed that this happens "through evaluation capacity building and nurturing an evaluation culture," (Douglah, Boyd & Gunderson, 2003, p. 7).

Labin, Duffy, Meyers, Wandersman & Lesesne (2012) defined evaluation culture as the "collective values, attitudes, goals, and practices that can support or hinder organizational change as related to evaluation," (p. 5). Gagnon, et al. (2018) concurred with other researchers (Cousins, 2004; Preskill & Boyle, 2008) that, when fully integrated into organizational culture, evaluative inquiry in the forms of both program evaluation and performance assessment is fundamental to the creation of new knowledge and capacity building.

Franz, Arnold and Baughman (2014) noted that "a strong evaluation culture has developed across Extension," (p. 3) but acknowledged that much of the evaluation being done and reported did not adequately show the impacts and/or public value of Extension programs. Berrio (1999) found that organizational culture may vary within sectors of an Extension organization and recommended that future study consider subcultures of the organization as they

relate to organizational learning. Mayne (2009) observed that, although organizations espouse the importance of an evaluative culture, they often do little to intentionally build and maintain a supportive culture. Further, he noted that, "efforts are typically put into building systems of measurement and reporting, and … enhancing the capacity of staff" (p. 4). Absence of a supportive evaluative culture limits an organization's ability to effectively do and use evaluation, and, currently, in the case of Extension, it limits effective impact evaluation practice.

The cultural component relates to the TPB (Ajzen, 1985) given that attitudes toward evaluation may impact evaluation culture and climate; in turn, evaluation culture and climate can impact behaviors of those in an organization (Ajzen & Fishbein, 2005; Labin, et al., 2012; Vengrin, 2015). Taut and Brauns (2003) proposed strategies for creating a more positive evaluation culture by addressing factors underlying individuals' resistance to evaluation including control, self-efficacy beliefs, personal cost-benefit analysis, and power issues.

Attitudes toward Evaluation

Wells-Marshall (2012) examined factors that influence an individual's attitudes and behaviors regarding systematic program evaluation within Extension. She found that the communication of information, attitudes, and behaviors relative to regarding program evaluation were predictors of an organizations' willingness to engage in or learn about evaluation (p. 2).

Vengrin (2016), in her study of Extension educators, found that attitude had a significant relationship to study participants' perception of the importance of specific evaluation competencies. She also found that attitudes were significant in their effect on subjective norms and on intentions to complete evaluation behaviors, which concurred with findings of previous researchers (Ajzen & Fishbein, 2005; Ajzen & Madden, 1986; Fazio, 1990).

Urban, Burgemaster, Archibald, and Byrne (2015) addressed attitudes toward evaluation in general in their study of the quality of program evaluation. They found that two factors correlated with higher or lower quality program evaluations—both optimism regarding evaluation and the valuing of evaluation.

Volkov (2011) noted that evaluative thinking is a capacity and important mindset that influences individuals' and an organization's readiness, willingness, and ability to engage in effective evaluation. This attitudinal factor can contribute to organizational learning and support the integration of a more positive habitual practice of evaluation within the organizational context (p. 38). Patton (2008) stressed the value of "evaluative thinking" in advancing rigorous, effective evaluation practice within an organization.

Kitinoja (1989) posited that pre-existing attitudes toward evaluation influence county Extension administrators' utilization of evaluation results. Further, she recommended that Extension educators be trained to focus more on *why* to do evaluation than *how* to do it in order to increase their readiness to do and use evaluation results (p. 12). An individual's attitude toward program evaluation behaviors is critical to understanding their intention to act and the resulting evaluation behaviors (Ajzen, 2002; Ajzen, 2011; Ajzen, & Fishbein, 2014). Urban, et al. (2015) found that, when the quality of educational programs evaluated by their rubric was compared to attitudes about evaluation, staff from higher scoring programs had more positive, optimistic attitudes and valued evaluation more than those associated with lower scoring programs. Further, both affective learning and emotions play a key role in transformative learning (p.189). Thus, when evaluation requires the learning of new evaluation strategies, skills, and knowledge—as in the case of transitioning to performing impact evaluation within an organization such as Extension—changes are a necessity.

Wells-Marshall (2012) addressed barriers and facilitators to systematic evaluation using the Readiness for Organizational Learning and Evaluation (ROLE) instrument developed by Preskill and Torres (2000). The ROLE instrument, "Is designed to help an organization determine its level of readiness for implementing organizational learning and evaluation practices and processes that support it" (Russ-Eft & Preskill, 2009, p. 498). Others have provided insight into barriers to fully embracing and integrating assessment into Extension work, including a lack of motivation driven by external demands for accountability, a perceived lack of time or skill to conduct assessment, and the perception of lack of support and direction regarding their role in evaluation of program impacts (Olson, Skuza, Blinn, 2007; Wells-Marshall, 2012; Vengrin, 2016). Arum and Roksa (2011) expressed skepticism that external accountability affects assessment of program impact. Carpenter and Stimpson (2007) postulated that, "since scholarship and research are frequently not familiar tasks, they are not considered to be as enjoyable or even as necessary, as, say, … planning a program, or any of the thousands of other tasks confronting busy Extension workers," (p. 272).

Aversion to being evaluated has long been an issue; increased attention to program evaluation may have exacerbated this over the past few decades (Bechar & Mero-Jaffee, 2014; Schwartz & Struhkamp, 2007). The concept of excessive evaluation anxiety, proposed by Donaldson (2007) and colleagues (Donaldson, Gooler, & Scriven, 2002), is an especially extreme, negative reaction and may manifest itself during any part of the evaluation process (American Psychiatric Association, 2016; Bechar & Mero-Jaffe, 2014; Taut & Brauns, 2003). Taut and Brauns (2003) noted that *reactance,* a strong aversive state brought on by a perceived threat to one's freedom of choice and decision making (Brehm and Brehm, 1981), results when evaluation is perceived as an exertion of power or as a control measure (Taut and Brauns, 2003).

Resistance to implementation of evaluation practices and a lack of openness to learning from experience can impede systematic evaluation efforts (Bechar & Mero-Jaffe, 2014). However, the implementation of evaluation efforts may increase the performance of an organization and serve as an agent of change in and of itself (Bechar & Mero-Jaffee, 2014; Patton, 2008).

Taut and Brauns (2003) suggested that these human factors can be influential when dealing with evaluation: a) trust in the evaluator, b) competence of the evaluator, c) prior evaluation-related learning experiences, and d) social skills of the evaluator. Perceptions of evaluation can have a direct impact on its use as well as on the actual behavior of evaluators or those being evaluated (Patton, 2008; Taut & Alkin, 2003). This may apply to impact evaluation as well as to evaluation in general; this study will provide insight into attitudes specifically regarding impact evaluation.

Attitude, Organizational Learning Culture, and Readiness for Change

Botcheva, White, and Huffman (2002) found that learning culture is an important factor for both the implementation of systematic outcome evaluation efforts and for success in obtaining external funding. They also suggested that systematic evaluation efforts can serve as a change agent for creating a culture that values learning within an organization. In their study, they utilized the Assessing Learning Culture Scale developed by Botcheva, White, and Huffman, (2001). Other researchers have noted that, among the relatively few existing measurement tools designed to assess evaluation capacity in organizations, most focus on cultural and structural aspects and rely heavily on psychometric, self-reported attitudes of professional staff (Botcheva, White, & Huffman, 2002; Suarez-Balcazar, et al., 2010; Volkov & King, 2007).

Researchers have posited that resistance to change is a factor in organizational learning and development (Patton, 2008; Preskill & Torres, 1999; Greene, 1988; Senge; 2006). Taut and

Brauns (2003) noted that various fundamental, psychological needs affect the level of resistance or reactance to evaluation. These include a) a basic *need for control,* b) a *need for positive self-concept,* and c) the *human tendency to avoid punishment and maximize rewards* (p. 259, 261). The researchers shared proposed strategies, based on interviews with administrators who lead and support evaluation in their organization that address the psychological needs identified. This theoretical focus on the human factors influential in carrying out evaluation provides support for user-oriented, participatory approaches to developing impact/outcomes evaluation capacity in organizations (p. 261).

Choi and Ruona (2011) challenged the importance of resistance to change and its influence on learning to perform evaluation in a position paper and literature review that focused on individual's readiness to change. They noted that individual readiness for change may reflect the concept of *unfreezing* proposed by Lewin (1947/1997b) and posited that understanding the conditions conducive to individual readiness for organizational change, instead of focusing on resistance to change, can support efforts to enact systematic outcome evaluation. The concept of individual readiness for change is a multifaceted construct involving individuals' belief in the appropriateness of the change, perception of the support by administration for the change, and personal cost-benefits of the change (Choi & Ruona, 2011; Holt, Armenakis, Feild, & Harris, 2007). Watkins and Marsick (1993) pointed out that an organization needs a "culture that is learning oriented, with beliefs, values, and policies that support learning" (p. 166). Other researchers have also stressed the importance of cultures of inquiry and generativity in facilitating organizational learning and change (Argyris & Schön, 1996; Senge, 1990).

Attention to educating employees to review their attitudes and related norms, stressing that the organization is a learning culture, and recognizing that employee contributions to the

process are valued can help employees focus on benefits rather than costs as they begin learning to plan and do impact evaluation (Choi & Ruona, 2011). These strategies prepare and support employees' readiness for change and learning. Bechar and Mero-Jaffe (2014) explored the complex associations among different factors that account for the successful implementation of outcomes evaluation. They found that a learning culture is foundational to implement research-based evaluation efforts necessary for outcomes focused evaluation. For impact evaluation and associated ECB work to be effective, individuals need both a clear understanding and a positive attitude toward doing evaluation and must value evaluation practice (Vengrin, 2016; Labin, 2014; Urban, et al., 2015; Clinton, 2014; Botcheva, White, & Huffman, 2002).

Little research on attitudes toward *impact* evaluation in Extension was found in the literature. Studies by Vengrin (2015), Workman (2010), and Wells-Marshall (2012) addressed attitudes about evaluation in general but did not specifically address impact evaluation.

Evaluation Competencies

Since it was legislated and founded, Extension has been called on to facilitate individuals' learning, business (primarily agricultural) successes, and to evaluate the effectiveness of educational programs (Wells-Marshall, 2012). Vengrin (2016) found these evaluator competencies to be most often mentioned as critical to Extension program impact evaluation: develops evaluation designs, specifies program theory, attends to issues of evaluation feasibility, shares evaluation expertise, and applies evaluation competencies to organization and program management.

Evaluation skills are regarded as core competencies and training priorities for Extension professionals (Arnold et al., 2008; Diem, 2009; Kluchinski, 2014; Lekies & Bennett, 2011; National Professional Development Task Force, 2004; Schwartz & Gibson, 2010). Vengrin

(2016) found that, while Extension educators reported being skilled in many evaluation competencies, there were gaps between what experts view as important and where Extension educators' skill sets lie.

Dozois, Langlois, and Blancher-Cohen (2010) and Volkov (2011) listed a range of competencies for developmental evaluators, including: caregiver, coach, strategist, facilitator, researcher, observer, advocate, and "tolerator of complexity and uncertainty" (p. 28).
King and Stevahn (2013) have been leaders in the recognition and promotion of the need for defined evaluation competencies. They shared the following tenets regarding evaluator competencies: a) technical knowledge and skills are the most important category of evaluator competencies; b) competencies that program evaluators need depend in large part on specific evaluation settings; and c) professional training in program evaluation is essential to build evaluator competencies (King & Stevahn, 2013). King and Stevahn also noted that some evaluation competencies may be intuitively discovered through evaluation practice in addition to gaining them through training and professional development.

New Extension practitioners identified evaluation and assessment competencies as areas that were underdeveloped in their preparation and/or daily practice (Franz, 2007). Even mid-level Extension educators – specialists often given assessment responsibilities – continue to struggle to gain confidence with respect to assessing program impacts (Vengrin, 2016).

A review of recent Extension literature provided ample evidence that program impact evaluation has been and continues to be an essential dimension of practice (Workman & Scheer, 2012). Extension administrators have been exploring ways to infuse better program impact evaluation and assessment into the ongoing practice of Extension education (Cook, 2016).

The challenge of building impact evaluation competencies is augmented since evaluation is not always perceived as essential. Vengrin (2016) found that Extension educators felt assessment competencies were crucial to the organizations' success yet indicated they did not highly value those competencies. The paradox that evaluation is clearly recognized as a critical competency, yet remains undervalued, is true for many Extension educators. Volkov (2011) noted there is "proverbial resistance" to evaluation and posited that resistance may be less an "anti-evaluation sentiment overall, but rather a rejection of bad evaluation and/or evaluators," (p.27).

Others provided insight into barriers to embracing and integrating impact assessment into Extension work, including a lack of motivation driven by external accountability demands, perceived lack of time or skill to needed to conduct assessment, and perceived lack of support and direction regarding their role in evaluation of program impacts (Wells-Marshall, 2012; Olson, Skuza, Blinn, 2007; Vengrin, 2016). Arum and Roksa (2011) expressed skepticism that external accountability affects assessment of program impact. Carpenter and Stimpson (2007) further postulated that, "since scholarship and research are frequently not familiar tasks, they are not considered to be as enjoyable or even as necessary, as, say, … planning a program, or any of the thousands of other tasks confronting busy Extension workers" (p. 272).

Negative, conflicting, and/or complacent attitudes of Extension practitioners further compound the challenge of developing impact evaluation competencies. Patton (2008) noted a conflict in how Extension professionals view themselves: as "educator" versus "change agent." This conflict still exists and affects how Extension professionals approach programming and impact evaluation. It appears that accountability pressures and the subsequent promotion of impact evaluation competencies in Extension have not motivated Extension professionals to fully embrace and integrate impact assessment into ongoing practice (Guion, et al., 2007). This

researcher concurs and views the challenge of infusing impact evaluation into Extension practice as especially critical at this juncture of funding cuts and conflicting public priorities.

A scholarship of inquiry is needed to enhance program impact evaluation in Extension. Faculty at state, regional and county levels must become more focused on learning to plan for and measure program impacts. Results of more rigorous impact-focused evaluation may be used to provide evidence of educational programs' private and public value, as well as to improve programs (Franz, 2009, 2011; Franz et al., 2014; French & Morse, 2015; Kalambokidis, 2011).

Self-assessment of Impact Evaluation Competencies

The benefits of self-assessment of competencies in general is to bring to the attention of the learner where they have strengths, where there are gaps in their knowledge, to raise awareness of unfamiliar information, and to guide action (Fetterman, Kaftarian, & Wandersma, 1996). Once individuals know what their strengths are, they may be able to share those strengths with coworkers and team members to put them to use (Brown & Dutton, 1995). Gaps in competency skills, knowledge, and attitudes may be addressed through training or staff development to increase competency and confidence in using them (Bandura, 1989). On the job training and experience is desired by many and seen as having some of the highest potential to allow for learning and application of that learning. Situated learning results in strong, constructed knowledge that is contextual (Lave & Wenger, 1991).

Research on individual competency building underscores the importance of mentoring, evaluation skill practice, and authentic evaluation use as crucial to ensure that Extension professionals move beyond knowledge acquisition to skill mastery in evaluation (Arnold, 2006; Baughman et al., 2010; Baughman, Boyd, & Franz, 2012; Dillman, 2013; McClure et al., 2012; Morford, Kozak, Suvedi, & Innes, 2006; Silliman, et al, 2016). This applies particularly in the

case of program impact evaluation, which is viewed as more complex due to its longitudinal, comprehensive, and change-focused nature.

Roles of Evaluators

Themessl-Huber, Harow and Laske (2005) posited that an evaluators' role is characterized by the competencies required, the functions performed, how they collaborate with stakeholders, and other contextual factors associated with a particular program and organization. Volkov (2011) outlined the following potential roles for internal evaluators: change agent, educator about evaluation, ECB practitioner, decision-making supporter, consultant, researcher, advocate and organizational learning supporter. Roles affect the competencies required to perform them.

Given the diverse roles of evaluators in Extension, different staff may need different competencies to do effective evaluation and, in today's context, to do effective *impact* evaluation to *provide evidence of change* resulting from Extension educational programs.

According to Stone and Bieber (1997) and Stone and Coppernoll (2004), competency development is a participatory process, and it provides Extension professionals with an opportunity to identify the knowledge, skills, and behaviors to obtain the best results as well as to identify the skills and functions that are no longer effective (p. 12).

Ghimire and Martin (2015) reported that Extension educators self-assessed needs for training in evaluation capacity building for the following competencies: "assessing impact of a program, developing and implementing surveys, analyzing and interpreting survey results, and using impact data for further planning," (p.11-12). Their findings have implications for designing and implementing impact evaluation competency development programs for professionals in all types of positions in Extension.

Internal Evaluator Roles

A "substantial share of evaluation work done nationally and worldwide is implemented internally" according to Volkov (2011, p. 25) and others (Fitzpatrick, Sanders, & Worthen, 2004; Love 1991, 2005; Sonnichsen, 2000). Few researchers have attempted to describe the range of roles performed by internal evaluators (Volkov, 2001; Love, 1991; Patton, 2008; Sonnichsen, 2000). Understanding the roles of internal evaluators is fundamental to ensure their effective and credible evaluation work (Volkov, 2011).

Scriven (1991) defined internal evaluation as work completed "by project staff, even if they are special evaluation staff—that is, even if they are external to the production/writing/ service part of the project" (p. 197). Volkov (2011) noted that internal evaluators play an important role in supporting an organization's program management and related decision making. Love (2005) asserted that internal evaluation is akin to applied research and supports organizational learning. Duffy (1994) observed that internal evaluation is similar to action research. Further, Volkov (2001) offered a systems type definition of internal evaluation:

> Internal evaluation is a very complex endeavor that involves important variables and conditions, including evaluator competencies, organizational culture and climate, stakeholder values, and external influences (p. 28).

Volkov (2011) noted that developing evaluation capacity in internal evaluators, such as those in Extension, benefits from the systematic promotion of positive change involving specific and general evaluation capacity building, supporting distributed decision making, promoting individual and organizational learning, and increasing evaluative thinking to support the role of the internal evaluator. Internal evaluator roles involve multiple facets and are complex.

Role Theory. Role theory literature presents conflicting research on whether filling multiple roles results in positive or negative effects on individuals' psychological well-being. Van der Horst (2016) noted that, overall, role theory research regarding the effects of filling multiple roles has shown that positive results are more likely than negative. Participating in multiple roles has been found to lead to role conflict, to role enhancement, to role overload, role confusion, role balance (Marks & MacDermid, 1996) and other outcomes. Marks (1977) posited that level of commitment affects and is affected by multiple role participation.

Evaluator Roles in Extension

Evaluators in Extension include various stakeholders that fund Extension, the different functions of staff (administrative, field staff, and faculty specialists), and other stakeholder partners in the private and publics sectors.

Scholars of evaluation practice have found that identifying "evaluation champions" benefits organizations as these individuals can serve as catalysts for handling accountability, evaluation learning and capacity building, and for innovation (King, 2007; Silliman, Crinion & Archibald, 2016a; 2016b). Further, evaluation champions advocate for the importance of program evaluation, model evaluation best practices, mentor their peers in the development of program evaluation skills, help shape evaluation-related policies, lead staff development and training, and help evaluate Extension ECB and evaluation practices. This has been the case in Extension (Silliman, et al., 2016a, 2016b; Taylor-Powell & Boyd, 2008) and in other organizations (King, 2007; Preskill & Boyle, 2008; Scheirer, 2005).

The role Extension evaluation specialists and others who have impact evaluation as a major part of their job responsibilities include engaging in evaluation capacity building—including mentoring and developing evaluation skills in Extension educators, serving on impact

evaluation teams, leading workshops in evaluation and impact evaluation—and assisting in creating a more supportive evaluation culture within the organization.

Larson (2016) noted that, in a relatively recent article, Wilcox and King (2013) explained that competence is an abstract concept that describes the state of being competent while competency is a more concrete concept that involves possessing knowledge, skills, and attitudes associated with being qualified. Thus, competency can be measured against accepted standards and improved by training and development (Larson, 2016; Parry, 1998). For this reason, Extension educators may benefit from reviewing and determining competencies which may help them be more effective in their program impact evaluation work.

Theoretical Foundations for Research

Various theoretical approaches may be useful in examining impact evaluation attitudes, behaviors, competencies, and related factors of Extension professionals come from social psychology. One is the MODE theory of attitude and behavior change (Fazio, 1990). A second is the Theory of Planned Behavior (Ajzen, 1980; 1991) which addresses not only attitude and behavior but allows for the inclusion of related demographic factors such as experience, related education and training. The theory of developmental evaluation (Patton, 2010) may also have applications for the study of ECB and program impact evaluation in Extension.

Luzar and Diagne (1999) noted that social psychology literature has documented the role of attitudes as reliable predictors of behavior (Ajzen, 1988; Ajzen & Fishbein, 1980; Fazio, Powell, & Williams, 1989). Mark, Donaldson, and Campbell (2011) observed that social psychology addresses many topics relevant to evaluation including processes affecting behavior change; effects of perceived, subjective norms; attitude change and persuasion; individual and collective decision making; and related processes.

MODE Attitude-to-Behavior Model

The MODE model describes processes through which attitudes can affect individuals' judgments and behavior (Fazio, 1990). The model focuses on two basic types of attitude-to-behavior processes – spontaneous processes versus deliberative processes. Attitudes are categorized as either implicit or explicit; attitudes interact with the processes that lead to behavior.

The MODE model considers *M*otivation and *O*pportunity as the key *DE*terminants of which type of process is most likely to operate—spontaneous or deliberate. Opportunity is seen as a "gateway" to facilitate processes that, in turn, influence judgments and behavior. Opportunity may include both the time and resources needed to allow a more deliberative process. Motivation may stem from an individual's desire for accuracy (Schuette & Fazio, 1995), a strong sense of accountability (Sanbonmatsu & Fazio, 1990), a concern with social desirability or norms (Fazio, Jackson, Dunton, & Williams, 1995), or other contextual factors.

The model includes many of the types of factors that can affect impact evaluation in Extension and could be used to explore attitudes (implicit/explicit, positive/negative), motivation (culture, reporting mandates), opportunity (resources and time available), and the types of processes that lead to effective impact evaluation behaviors.

Theory of Planned Behavior

The Theory of Planned Behavior was developed by Ajzen (1991), who found that the key factors of *attitude*, *subjective norms* and *perceived behavioral control* can impact an individual's intention to complete a behavior (1991). Intentions, coupled with perceived behavioral control, were found to account for variance in the behavior and to assist in predicting of actual behavior (Ajzen, 1985; Armitage & Conner, 2001; Orbeil, Hodgkins, & Sheeran, 1997).

The three major factors of attitude, subjective norm, and perceived behavioral control, each play a significant role in explaining and predicting behaviors (Ajzen, 1991, 2002, 2003). Figure 2-7 illustrates the theory's basic concepts and their relationships.

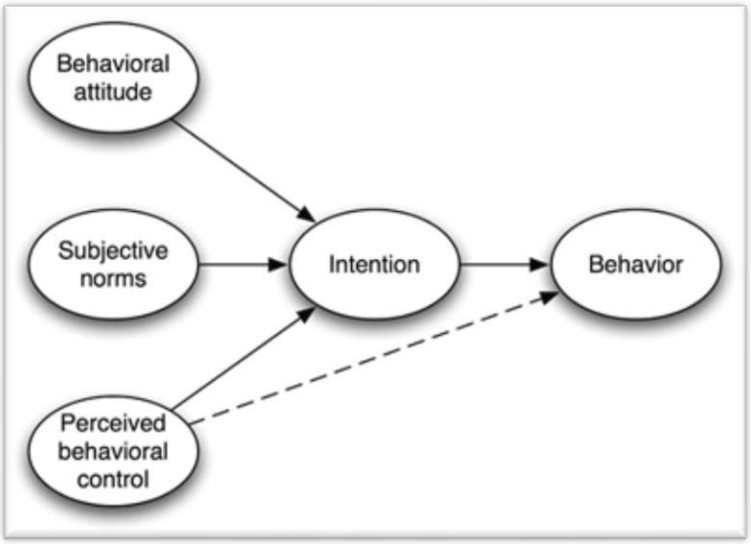

Figure 2-7. Basic Theory of Planned Behavior model (Ajzen, 1991).

For some scenarios, utilizing this theory can help to guide the design of interventions that may stimulate or change behavior (Gargani, 2012; Daigle et al., 2002; Greaves, Zibarras, & Stride, 2013; Lin, Chan, & Wei, 2006; Meng, Othman, D'Silva, & Omar, 2014). In order to catalyze behavioral change, understanding each component of a decision to complete a behavior is crucial (Vengrin, 2016).

The Theory of Planned Behavior was presented as a well-developed theory describing a complete process in a recent article on key theories useful for studying social-ecological systems (Schlüter, Baeza, Dressler, Frank, Groeneveld, Jager, Janssen, McAllister, Müller, Schwarz, & Wijermans, 2017). The theory includes factors ranging from the formation of beliefs and attitudes, to normative beliefs, to intentions to perform a behavior, and, finally, to actual execution of the behavior based on those beliefs and attitudes (p. 23). Given that attitude

contributes to culture, this theory is adaptable to inclusion of the factor of culture for research (Ajzen & Madden, 1986; Labin et al., 2012). Perceived behavioral control refers to "beliefs regarding the possession of requisite resources and opportunities for performing a given behavior" (Madden, Ellen, & Ajzen, 1992, p. 4). The theory also allows the inclusion of additional explanatory variables such as demographics.

The theory explains behaviors, the intention to perform them, and factors involved in the behavioral decision-making process (Ajzen, 1991; Gargani, 2012). Han (2015) observed that the efficacy of the TPB in explaining many types of behaviors has been validated in diverse situations (Ajzen & Driver, 1991; Han & Kim, 2010; Lam & Hsu, 2004, 2006; Oh & Hsu, 2001). Gargani (2012), Vengrin (2016), and Han (2015) asserted that the TPB (Ajzen, 1985; Ajzen, 1991) is superior to theories such as the MODE model, the reasoned action theory (Fishbein & Ajzen, 1975), Vroom's (1964) expectancy theory, and others for the examination of human behavior and the many factors that influence behavior within a particular cultural context.

Attitude is a key factor in the Theory of Planned Behavior and is reflected in behavioral intention (Ajzen, 1991, 2003; Conner & Armitage, 1998). As with motivation, if an individual does not have a positive attitude toward a behavior it is unlikely that they will have a strong intention to perform the behavior (Ajzen, 2006; Conner & Armitage, 1998; Fazio, 1990). External motivating factors, such as demands from leadership may impact motivation and intention to perform a behavior, especially if an individual does not find it desirable (Vengrin, 2016). Attitude to perform a behavior can be offset by other factors and result in the behavior being carried out even when these external motivators exist.

Subjective norms are social factors related to how individuals perceive the social desirability of a behavior. This can generate external pressure, either authentic or imagined, to

perform or avoid the behavior. Subjective norms may be considered a less influential factor in behavior prediction, because an individual's attitude regarding the behavior can overshadow the social pressures (Ajzen, 1991; Armitage & Conner, 2001). These factors may, however, still play a role in the intention and completion of a desired behavior.

In order to predict behavior by utilizing the Theory of Planned Behavior, several conditions must be met (Ajzen, 1991). Perceived behavior control must correspond directly to the behavior to be predicted, intentions and perception must remain stable between the time of assessment and observation of the actual behavior, and perceived behavioral control must be accurately reported or predicted (Ajzen, 1991). Depending on the behavior, these factors can vary in their ability to predict behavior. Intention and perceived behavioral control interact in the prediction of behavior (Ajzen, 1991).

Behavioral intention is defined as a person's perceived likelihood or "subjective probability that he or she will engage in a given behavior" (Committee on Communication for Behavior Change in the 21st Century, 2002, p. 31). Given that TPB asserts that intention strongly correlates to actual performance of a behavior, this approach has much to offer the study of impact evaluation behaviors in Extension. Per TPB, intention is guided by *attitude* toward evaluation behaviors, *subjective norm* (culture and other perceived influences); and *perceived behavioral control* (self-efficacy and perceived ease or difficulty based on both previous experience and anticipated barriers) (Ajzen, 1991).

Perceived behavioral control may also be influenced by information and opinions obtained from others who have experience regarding the target behavior (Ajzen, 1991). This can shape how individuals view the behavior and their self-efficacy regarding it. If an individual perceives that they have a considerable amount of knowledge and resources related to the

behavior, they are likely to feel more able to perform the behavior, thus exhibiting an increase in their perceived behavioral control (Vengrin, 2016).

The Confluence of Evaluation Capacity Development, Organizational Evaluation Culture, and the Theory of Planned Behavior in Extension Impact Evaluation

Vengrin (2016) chose to study evaluation behaviors, competencies, and related factors in Extension based on the intersection of the Theory of Planned Behavior and Evaluation capacity Building. This study uses a similar theoretical approach; however, it adds individual and organizational demographic factors found in literature and focuses specifically on *impact* evaluation behaviors in Extension. This descriptive, correlational study of Extension educators' impact evaluation attitudes, skills, and behaviors utilizes the Theory of Planned Behavior, organizational learning theory regarding evaluation culture, and competency and capacity development theory as a multi-faceted lens through which to view the relationships among factors that influence impact evaluation behaviors and practice in the Extension organization.

While many have addressed the issue of negative attitudes toward evaluation, in general, few have suggested strategies for changing those attitudes (Bechar & Mero-Jaffe, 2014; Kitinoja, 1989; Patton, 2008; Taut & Brauns, 2003). Research on an organizations' learning culture and individuals' readiness for change as a way to focus learning about and valuing evaluation within an organization offers an approach to improving impact evaluation culture within Extension (Botcheva, White, & Huffman, 2002; Choi and Ruona, 2011; Taut & Brauns, 2003).

Research by Baum (2015) illustrated how complex interaction among factors may influence attitudes and beliefs regarding assessment and evaluation. Baum included the factors of self-efficacy, professional standards, competency, reflection on the purposes of evaluation, accountability and institutional support. Baum recommended that future research consider the

intersection of the individual and organizational environments to better understand the relationships among them.

For these reasons, this study combined several theories to explore potential relationships among the constructs and factors which may influence Extension impact evaluation.

Summary

The need for Extension professionals to provide evidence of program impacts is greater today than ever before due to increased accountability and pressure to demonstrate relevance in an era of possible declines in funding. Impact documentation needs to stem from Federal compliance requirements and from a need for publicly available outcomes data for institutional and state comparisons. Extension must demonstrate a positive return-on investment to clientele and the public by documenting evidence of impacts in terms of change in social conditions, considered a long-term outcome. To move evaluation practice to address a higher level of evidence of program impact, Extension must focus on reporting impact outcomes.

Program evaluation studies reported in the *Journal of Extension* during the past three decades showed some evidence of effectiveness but did not address evidence of "program impact" specifically. Some progress has been made but Extension educator/evaluators continue to report the need for more skills in impact evaluation. A key issue is that, while complex Extension educational programs and interventions can be challenging and expensive to evaluate, there is a critical need to ensure that Extension professionals possess the essential competencies needed to conduct high quality and technically sound impact evaluations. Program impact evaluation in Extension increasingly involves more team approaches, such as interdisciplinary teams at the state, regional, and national levels.

An emerging focus within Extension is the impact of programs in creating public value – including the value to those who do not participate in their programs (Kalambokidis, 2004). Before Extension can document programs' public value, they must continue to produce strong research on Extension's private value to those who are impacted directly by programming.

Behavioral research has shown that an individual's attitudes, beliefs, and expectations are important predictors of individual behavior (Ajzen & Fishbein, 1980; Bandura, 1977b). Demographic factors such as experience, staff development and training, academic background, and other factors may also influence both individuals' intent to perform a behavior and their actual performance of that target behavior.

Larson (2016) noted that there is a gap in research about evaluation competency development; thus, it is prudent to pursue inquiry that explores evaluation competencies needed to do rigorous and effective impact evaluation in Extension.

This study addressed the following factors that affect the accomplishment of impact evaluation in Extension: competency levels and perceived value of those competencies to Extension educators; the organizational culture and context for learning to plan, do and utilize impact evaluation results; the attitudes, beliefs, and perceived control over impact evaluation behaviors; and the relationships among the many factors, including demographics factors such as individuals' previous experience with impact evaluation, years in Extension, program area, and type of position within the organization and organizational factors such as teamwork, program planning models used, and support for impact evaluation. Viewing the issue of completing effective impact evaluation in Extension through a lens of complex, multi-faceted factors will serve to illuminate potential relationships that affect educators' impact evaluation behaviors within the organization.

CHAPTER 3. RESEARCH METHODOLOGY

In an era of increased accountability and funding cuts, Extension and other publicly funded organizations must demonstrate relevance, impact, and return-on-investment regarding their educational programs (Kluchinkski, D., 2014; Decker, 1990; Kelsey, 2008; Rennekemp & Engle, 2008). Impact evaluations must address Extension educational programs' potential public value and impact on society (Kalambokidis, 2017). Impact is defined as "the reportable, quantifiable difference or potential difference a program makes in the lives of real people. It shows a sustainable societal, environmental, and/or economic change," (eXtension Impact Online course, 2017).

Organizations that require impact evaluation can better support effective impact evaluation by increasing their understanding of important evaluation competencies within the organization; evaluation attitudes and behaviors of personnel; and their organizational evaluation culture. Gaining insight into these interrelated individual and organizational factors can help Extension administrators to more effectively convey the need for impact evaluation to staff and to design ECB efforts to promote higher quality and better focused evaluations of program impacts and outcomes.

The purpose of this study was to employ a comprehensive, systems approach to identify potential relationships among the individual and organizational factors that influence the practice of impact evaluation (IE) in Cooperative Extension. Factors included were a) perceived skill levels and importance of impact evaluation competencies; b) impact evaluation culture of the organization; c) impact evaluation behaviors and their antecedents; and d) individual and organizational demographic factors such as educational level, program area, years of Extension experience, training, program planning models used, amount of teamwork, and more.

The intent of this chapter is to describe the methodology used to examine the impact evaluation related relationships described above. The descriptions that follow include the

research design; study population and sample; instrumentation overview; data collection strategies and methods; and the approaches to data analysis.

Research Design

This research was a quantitative, non-experimental correlational study of the attitudes, evaluation competencies, evaluation culture, evaluation behaviors, and individual and organizational demographic factors associated with Extension educators' impact evaluation practices in the North Central Region of the Cooperative Extension service.

The constructs for this research were chosen from the theoretical framework of the Theory of Planned Behavior (TPB) found in the literature and influenced by research relating to organizational evaluation culture, evaluation competency requirements, impact evaluation, and related individual and organizational demographic characteristics.

Research methods employed included administration of an online, quantitative survey to assess the evaluation competencies valued by Extension educator/evaluators as important for their practice of impact evaluation and at what level they feel confident that they can perform those competencies. The study also employed quantitative survey methodology to examine the evaluation culture and evaluation behaviors within Extension. Two open-ended questions provided qualitative responses to help enrich understanding of the quantitative data.

Online survey research was used and appropriate because data had to be obtained from individuals in geographically dispersed locations in an inexpensive and quick manner (Creswell, 2008, 2005; Fowler, 2014; Russ-Eft & Preskill, 2009; Wells-Marshall, 2012). Survey research has also been shown to be a reliable and valid means to gather information about characteristics, attitudes, or behaviors of a population (Wells-Marshall, 2012).

The research design is congruent with the theory-based complex, systems approach to exploring possible relationships among the parameters deemed to be involved in the performance of behaviors associated with impact evaluation in Extension. The research goal was to provide a more comprehensive understanding of the antecedents and interrelationships among the constructs included in the study. A common practice in research involving the TPB is to determine its usefulness in combination with other models and frameworks.

A major feature of this research design was the decision to use Structural Equation Modeling (SEM) statistical analysis technique. The use of SEM to organize and implement this study facilitated simultaneous exploration of many constructs theorized to be involved in impact evaluation behaviors in the context of Extension educational program evaluation. Furthermore, SEM, similar to regression, is appropriately used for analysis when a research design is non-experimental (ex post facto) (Frey, 2018; Kline, 2016).

Kline (2016) noted that SEM can be used to represent and test quantitative models constructed based on a substantive theoretical framework. Further, benefits of employing SEM analyses include: a) one can examine complex relationships among variables; b) it allows for simultaneous parameter estimation; c) models can incorporate and allow for measurement error; d) it can include both latent and observed variables; and e) it can account for variables which may have been omitted from a model (Kline, 2016).

Quantitative data to represent the constructs included in the theoretical model for the study were collected using three Likert-scaled survey subscales. A self-reported online survey was administered to study participants. The survey was designed to measure individuals' perceptions of the importance of a list of competencies needed to do effective impact evaluation, and their rating of their skill level for each of those competencies; their impact evaluation

behaviors; and the impact evaluation culture within their Extension organization. Demographic data including years of employment in Extension, training in evaluation, program content/ discipline area, highest academic degree level, state (location), teamwork experience, and previous evaluation experience were collected. Data analysis was planned to examine the relationships among those factors to identify potential relationships and antecedents of impact evaluation behaviors. Qualitative data were also collected via two open-ended items to add depth and richness to the data. That data was examined thematically and correlated with other factors to identify potential complementary or conflicting relationships among the data.

Specification of the Structural Equation Model

Specification of an initial theoretical model to be used to frame the potential relationships among the many factors that may influence program impact evaluation behaviors in Extension was necessary prior to data collection in order to use SEM for data analysis.

For this study, model specification began with the creation of a preliminary theoretical model including the basic constructs in the Theory of Planned Behavior (TPB) as shown in Figure 3-1 in blue and listed in Table 3-1. The TPB constructs were measured using Likert-scaled items as indicators of each construct. Composite means for each construct were calculated to provide a value for the variable to be used in the SEM analysis (see Appendices G and H).

The constructs of culture, competency by perceived skill level, and competency by perceived level of importance were added to the model next based on capacity building literature, organizational and evaluation culture literature, and learning organization theory. The competency by perceived skill and perceived importance were measured using Likert scales; composite means were calculated for each variable to provide values to be used in the SEM analysis of the Actual Initial Model following respecification based on methodological and data.

The constructs added for TPB, impact evaluation culture, and competency by perceived skill level and perceived importance made up the expanded TPB model. The constructs in this core section of the model are all endogenous variables. These endogenous variables are defined in Table 3-1 and shown in Figure 3-1. Endogenous variables in a SEM model may be independent and/or dependent variables. There is one dependent endogenous variable in this study—the desired target behaviors of actual impact evaluation (IE) behaviors.

Table 3-1

Endogenous Variables in the A Priori Theoretical Model

Variable/Construct	Definition	Source(s)	Exogenous/ Endogenous
Actual Impact Evaluation (IE) Behaviors	Types of IE behaviors performed by Extension professionals.	Roche, 1999.	Endogenous
Behavioral Intent	Motivational factors that influence behavior; indications of how much of an effort they are planning to exert, in order to perform the behavior.	Ajzen, 1991.	Endogenous
Attitude Toward IE	A disposition, based on beliefs, to respond favorably or unfavorably to a specific behavior (impact evaluation); attitude is expected to predict and explain human behavior.	Ajzen & Fishbein, 1999.	Endogenous
Subjective Norm	Social factors regarding how an individual perceives the social desirability of a behavior within the context where the behavior is expected to be performed.	Ajzen & Fishbein, 1975.	Endogenous
Perceived Behavior Control (PBC)	PBC is essentially the same idea as self-efficacy; the measure of a person's perception of his or her ability to complete a task.	Ajzen & Fishbein, 1999.	Endogenous
Culture (IE focused)	Organizational evaluation culture regarding impact evaluation.	Preskill & Torres, 2001.	Endogenous
Competency by Skill	Perceived level of skill for IE competencies self-rated by study participants.	American Evaluation Association (AEA), 2018	Endogenous
Competency by Importance	Perceived importance of IE competencies self-rated by study participants.	AEA, 2018	Endogenous

Note. Variables in blue are factors in the Theory of Planned Behavior (Ajzen, 1991; Ajzen & Fishbein, 1975, 1999); variables in green are factors from the Vengrin Extension Evaluation model (Vengrin, 2016).

Exogenous variables made up the remainder of the specified SEM model. These variables are shown in Tables 3-2 which outlines individual factors and Table 3-3 which outlines organizational factors. In the graphical concept model show in Figure 3-1 the exogenous variables are found in the far left-hand section of the model. The exogenous variables are factors that have been found by previous researchers to influence evaluation behaviors in general in Extension (Lamm, 2012; Wells-Marshall, 2012; and Vengrin, 2016).

Table 3-2

Exogenous Individual Variables in the A Priori Theoretical Model

Variable/Construct	Definition	Source(s)	Exogenous/ Endogenous
Experience (yrs)	Years employed in Extension.	Vengrin, 2016.	Exogenous
Program Area	Primary program area responsibility: Agriculture/Natural Resources (ANR), 4-H Youth Development, Family & Consumer Science (FCS), Community Resource Development (CRD), Other.	Vengrin, 2016.	Exogenous
Position Type	Position type: County Agent / Ext. Educator/Program Coordinator, Program Assistant (e.g. SnapEd, 4-H), Specialist, Administrator (state, dept., district), Other.	Wells-Marshall, 2012; Ghimire & Martin, 2013,	Exogenous
Degree Level	Highest academic degree: PhD/EdD, Masters, Bachelors, Associate Degree, Highschool, Other.	U.S. Census Bureau, 2015.	Exogenous
Training in IE	Training in impact evaluation (Y/N or Maybe)	Vengrin, 2016	Exogenous
Training Type(s)	Type(s) of training in Impact evaluation (On-the-Job experience; Staff development, Face-to-face; Staff development, Online; Mentoring by administrator or colleague; Part of academic course; Grant-required; Self-directed study).	Ghimire & Martin, 2013; Wells-Marshall, 2012.	Exogenous
Roles in IE	Number of role(s) performed in impact evaluation – design, data collection,	Van der Horst, 2016.	Exogenous

Foundational to the development of impact evaluation capacity is a positive learning organization culture with respect to evaluation and, specifically, toward impact evaluation. Klein (2009) noted that an organization learns through its individual members and is, thus, influenced by individual learning. Argyris and Schön's theory of organizational learning posits that organizational learning is accomplished through individuals whose actions are guided by shared models (1974). Developing professional competence in impact evaluation requires intentional capacity building to become more effective at both the individual and organizational levels. This theory-based stance underlies the specified a priori theoretical concept model created to guide the SEM approach to this research. The model is shown in Figure 3-1.

Table 3-3

Exogenous Organizational Variables in the A Priori Theoretical Model

Variable/Factor	Definition	Source(s)	Exogenous/ Endogenous
Support	Separate evaluation/program planning unit in your organization?	Baum, 2015.	Exogenous
Teamwork	Proportion of impact evaluation work done through teamwork.	Berrio, 1999.	Exogenous
Location (state)	States in the North Central Region of Cooperative Extension.	North Central Region, 2018	Exogenous
Program Planning Model	Types of program planning models used in Extension: Kellogg Logic Model, Bennett's hierarchy, Kirkpatrick's model, and other.	Kellogg Foundation, 2004; Kirkpatrick, 1959; Bennett, 1875.	Exogenous

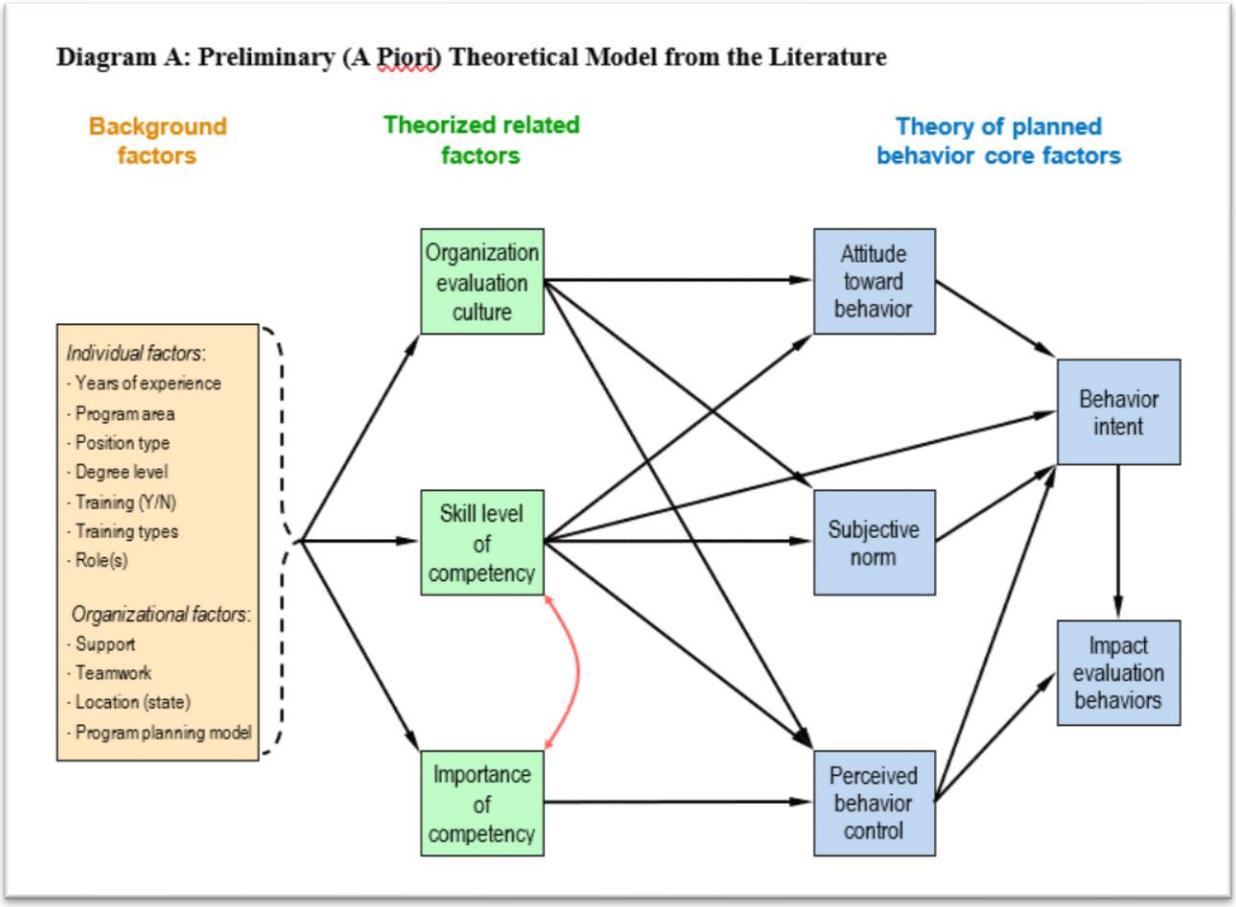

Figure 3-1. Theoretical model developed from the literature. Black arrows represent direct effects (pathways). The red double-headed arrow indicates correlated disturbance (error) terms.

Study Population and Sampling Frame

The population for the study were Extension educators at all levels of the organization who have responsibilities for impact evaluation of one or more educational programs and work in states included in the North Central region of Extension.

The sampling frame was the current list of Extension educator/evaluators provided by liaisons in each state's Extension Service. Non-probability sampling was employed to optimize contacting the maximum number of individuals, given the constraints of population access, including geographic dispersal. Study participants were recruited by working with evaluation leaders from each state within the North Central Region.

The issue of ensuring that a sample adequately represents the characteristics of the population from which the sample is drawn is essential to a conduct a reliable research study. An appropriate sample size must also be secured. Fowler (2009) advised that research study sample size must be congruent with the data analyses planned for the study. Therefore, guidelines for sample size determination were based on the Structural Equation Modeling (SEM) statistical analysis technique chosen for this study. To obtain reliable results when using SEM, researchers agree that a sample size of 200 is adequate to test a model (Buhi, Goodson, & Neilands, 2007; Kenny, 2014; Kline, 2011). Based on this recommendation and input from this study's SEM methodologist (Kline, 2016), the researcher's goal was to obtain 200-250 responses to facilitate a strong structural equation modeling analysis.

Measuring Constructs, Factors and Variables

The a priori theoretical conceptual model included the constructs and factors and the potential relationships among those factors. These constructs and factors were well-defined and the instrumentation was designed to measure each to ensure that the resultant variables were useable as legitimate factors in the model. Specification of the model is a key step in the structural equation modeling (SEM) process and was described earlier in this chapter.

The model contained one dependent variable, actual impact evaluation behavior, which was an endogenous factor in the model. Measuring actual impact evaluation behavior (actbx) was accomplished by asking respondents to provide information on their actual performance of a list of impact evaluation methods; a composite score of their self-reported level of performing the impact evaluation behaviors was calculated for use as the variable in the SEM analyses.

Other endogenous, independent variables in the Theory of Planned Behavior section of the model were the following constructs measured using Likert-scaled items to assess the factors

and provide values for use in the SEM analyses: attitude, subjective norm, perceived behavior control, and behavioral intent. Additional endogenous variables from related theory included competencies and organizational evaluation culture.

Exogenous variables were identified in the model as demographic factors including characteristics of both individuals and the organization. Individual factors were years of experience in Extension, program area (Agriculture, 4-H, family and consumer science, etc.), type of position (county-based, administrator, specialist), highest degree, role(s) in impact evaluation, training in impact evaluation, and types of training experience. Organizational demographic factors related to impact evaluation behaviors were support for evaluation, teamwork, program planning model(s) used, and location (state).

The constructs and observed, manifest variables included in the preliminary theoretical model provided clear guidelines for creation of the data collection instrument.

Instrumentation

To measure the constructs included in the conceptual model required a variety of measurement approaches in the data collection instrument. The instrument was adapted from Vengrin's instrument, the Evaluation Competencies, Culture and Behavior (ECCB) survey (Vengrin, 2016). The instrument also included an adapted version of Preskill & Torres' (1999) subscale on culture to assess impact evaluation culture specifically. Permission to adapt each instrument was obtained from the authors. Preskill & Torres' instrument demonstrated high reliability and construct validity in several studies. Vengrin's ECCB instrument was also found to demonstrate high reliability and validity in her study (2016). The ECCB instrument was revised by this researcher (Flack) and updated to reflect changes in the professional evaluation community in the United States and current impact evaluation priorities in Extension. Employing

a revised version of Vengrin's ECCB instrument which was designed for and used within the context of the Extension organization made it a very good fit to explore the relationships among factors that may affect impact evaluation behaviors in the target population for this study.

Three data collection subscales and demographic items were created to examine the constructs and variables determined to be important to impact evaluation behaviors within Extension. The subscales addressed a) evaluator competencies – perceived importance and skill levels; b) impact evaluation culture; and c) impact evaluation behaviors. Demographic items addressed the individual and organizational factors included as exogenous factors in the conceptual model. A copy of the instrument is provided in Appendix A of this document.

Subscale One – Competencies

Evaluation competencies were examined using a scale comprised of a list of evaluation competencies created by the researcher and reviewed for face validity by impact evaluation experts in Extension. The list of competencies included in the survey instrument for this study is a composite, edited list, based on the current list of Evaluator Competencies by the American Evaluation Association (AEA, 2018); Canadian evaluator competencies used by Vengrin (2016); and lists of evaluation competencies identified by Extension researchers including Ghimire and Martin (2013), Kluchinski, (2014); McClure, et al. (2012), and Wells-Marshall (2012).

The AEA competencies finalized in mid-2018 includes items in five domain categories of expertise and attitudinal factors important to the design, implementation, and use of high quality, effective evaluations. The list of impact evaluation competencies compiled for this study included items from each of the five AEA domains of a) professional practice, b) methodology, c) context, d) planning and management, and e) interpersonal.

Items for this research were adapted to represent competencies relevant to impact evaluation in Extension. Items in the final draft of the list were compared to those used in previous Extension evaluation competency studies to ensure that key competencies were not omitted (Ghimire & Martin, 2013; Kluchinski, 2014; McClure, et al., 2012; Urban, Burgemaster, Archibald, & Byrne, A., 2015; Vengrin, 2016; Wells-Marshall, 2012). Extension evaluation experts reviewed the list of competencies prior to finalization of the instrument.

Subscale Two – Impact Evaluation Behaviors

The third subscale of the survey addressed factors relevant to Extension impact evaluation behaviors. For this section, this researcher used guidelines for creating scales based on the Theory of Planned Behavior proposed by Francis (2004) and Ajzen (2011) to review, revise, and expand Vengrin's ECCB behavioral subscale. Vengrin's subscale included 16 items which were validated and tested for internal reliability in her 2016 study. Her behavior subscale was validated to measure the following four factors in the Theory of Planned Behavior as it applies to evaluation: a) attitude towards evaluation, b) behavioral intentions, c) subjective norms, and d) perceived behavioral control. Vengrin's validated items were adapted to address impact evaluation and included for this study.

The behavioral subscale created for the research instrument for this study contained additional items to further address attitudes toward *impact evaluation*, specifically, and to document actual impact evaluation behaviors. To increase construct validity, guidelines for conducting TPB studies and constructing a TPB questionnaire by Ajzen (2005) were followed. These items adapted from the guidelines were also edited based on items validated by Smith (2015) and Vengrin (2016): "It is entirely up to me whether or not I complete program impact evaluations this year," "My efforts to complete program impact evaluations are valuable," and "I

will make an effort to complete program impact evaluations this year." Two items adapted from examples by Ajzen and Fishbein (2005) for this study have previously been adapted, used, and validated by others (Montañ & Kasprzyk, 2015; Smith, 2015; Teo & Lee. 2010; Vengrin, 2016) to measure both perceived behavioral control (PBC) and subjective norm (SN) . This researcher's adapted versions of the items are: "Most people who are important to me professionally think I should use impact evaluation strategies in my work," and "If I wanted to, I could easily use impact evaluation competencies in my work to complete effective impact evaluation." Similar items employed by Smith (2015) had alpha coefficients of .91 and .83 respectively. Two other items adapted for this instrument, "I intend to use impact evaluation competencies and strategies in my work as an Extension professional," and "Using impact evaluation in my work is a good idea," were adapted from an example by Ajzen and Fishbein (2005). Sadaf et al. (2012) reported Cronbach's alpha values from .83 to .93 (p. 180) on similarly adapted items. Alpha values were calculated for this study instrument after data collection and are reported later in this chapter.

Subscale Three – Impact Evaluation Culture

The culture subscale created and validated by Preskill and Torres (2000) for their readiness for organizational learning and evaluation (ROLE) survey was utilized to further explore attitudes, subjective norms and administrative support. Instructions for completing this section of the survey directed study participants to answer with respect to the *impact* evaluation culture in Extension. The ROLE instrument was designed, "to help an organization determine its level of readiness for implementing organizational learning and evaluation practices and processes that support it" (Russ-Eft & Preskill, 2009, p. 498). Preskill and Torres (2000) found the ROLE instrument to be internally consistent (Cronbach's alpha was .97 across the Likert scale items for the entire instrument and a range of .83 to .94 for the subscales, including .92c for

Culture). The instrument's construct validity was established earlier (Preskill, Torres, & Martinez-Papponi, 1999). The ROLE culture subscale includes items in three sub-construct areas: a) Collaboration and Problem Solving – 12 items; b) Risk Taking – 5 items; and c) Participatory Decision Making – 10 items.

The ROLE scale, or portions of it, was previously used by two researchers to assess evaluation culture and readiness for learning regarding evaluation in general in Extension (Wells-Marshall, 2012; Vengrin, 2016). This study is the first to use the instrument to assess culture and readiness to do impact evaluation in Extension.

Demographics

Demographic items measured both individual and organizational factors included in the conceptual model. Individual factors included years of experience in Extension, program area, type of position, highest degree, role(s) in impact evaluation, training in impact evaluation, and types of training experiences. Organizational demographic factors related to impact evaluation behaviors were support for evaluation, teamwork, program planning model(s) used, and location (state). Age and gender were not collected to help ensure confidentiality.

Survey Item Formatting

The survey instrument included only positively worded items as recommended by Peterson et al. (2006). Using both positive and negative items in an instrument could have a detrimental effect on psychometric properties associated with the measures (DeVellis, 2016). Most Likert items were scaled using 6-point and 4-point scales to attempt to alleviate central tendency bias; the previously validated ROLE culture subscale, however, used a 5-point scale. The online survey was created using the Qualtrics Research Suite software.

The survey was designed to be completed by participants within 20 to 30 minutes. The survey instrument contains 42 competency items, each on two scales (importance and level of expertise); 27 items pertaining to evaluation culture (Preskill & Boyle's culture subscale); 25 items pertaining to Extension variables related to TPB behavioral factors; and 7 demographic variables, for a total of 143 items. The items were scaled according to literature guidelines and grouped in sections for logical completion by participants.

Protection of Human Subjects

The North Dakota State University Institutional Review Board (IRB) policies and procedures were followed throughout the duration of the research study. Approval was obtained prior to initiation of expert review of the instrument. No substantive changes were made to the instrument following review; thus, resubmission of the final survey instrument was not required. The NDSU IRB approval of the study served as the human research approval for all states involved in the study given that data were collected online and state liaisons served only to recruit participants. Copies of the approved IRB protocol were shared with liaisons in participating states when requested. A copy of the IRB approval notification is provided in Appendix D of this document.

Preliminary Work: Expert review of Instrument

Input on revisions of survey items to facilitate best data analysis were sought from research committee members and experienced Extension evaluation experts prior to finalizing the survey. Extension expert evaluators from states in the North Central Region of Extension were identified using a snowball sampling technique employing referrals by current evaluation leaders within their organizations who had evaluation as a primary job responsibility. Twelve individuals with expertise in Extension evaluation and specifically, in impact evaluation, were

invited to participate in the review; ten individuals accepted the invitation. Feedback from the expert evaluators was used to make minor revisions to the online data collection instrument. Feedback was used to clarify, delete, and/or add items in order to ensure that the instrument addressed each of the following constructs to be measured: impact evaluation culture, impact evaluation competency levels, importance of impact evaluation competencies, impact evaluation behaviors, and relevant individual and organizational demographic factors.

Data Collection and Analysis

Following dissertation committee approval of the study, Institutional Review Board (IRB) approval was procured to ensure that proper human subjects research protocol was followed throughout the process of conducting the research study. Data collection and handling protocols were outlined in the IRB application.

Data Collection Procedures

Following expert review and finalization of the online survey, Extension evaluation specialists and experts in states within the North Central Region of Extension were contacted to request participation in the study. The process to recruit participation of state Extension organizations in the study and liaisons in each of those states to assist with recruitment of individuals participants was done in collaboration with one of the researcher's supervisory committee members who was part of a regional group of evaluation experts in the North Central Region of Extension. The researcher drafted and share background information about the purpose and significance of the impact evaluation study. The committee member drafted an email which was then sent to her evaluation colleagues in Extension organizations in the North Central Region. The email introduced the researcher, the proposed study, and gave information on how to contact the researcher to indicate interest in participating in the study. A total of six

states within the region agreed to participate. Following notification from the Extension evaluation liaison that their colleagues had consented to support their state's participation in the study, the researcher scheduled phone conferences with each state liaison to review data collection procedures. Liaisons were colleagues familiar to Extension professionals in their respective states. Working with evaluation experts in each state increased the possibility that respondents would understand the importance of the research to their own evaluation efforts.

The choice to involve state liaisons in the recruitment of participants was made to increase the likelihood of better response rates. The initial invitation email with an embedded survey link was sent to each liaison. Liaisons then sent the emails, and two reminders, to all Extension professionals responsible for doing program impact in their states. State liaisons from five of the six states chose to forward the email to those on their staff while one gave permission for the researcher to use their email list to send the invitation to participate directly to staff using the Qualtrics program. Both protocols for distributing the initial invitations to participate and two reminder emails were approved in the NDSU IRB application.

The researcher conscientiously adhered to established guidelines for conducting educational research. Information was provided to potential participants on the first page of the online survey about confidentiality, anonymity, risks and benefits, and the informed consent required prior to being given access to the survey. The researcher provided contact information for questions and/or concerns about the research. Participants were also informed that study results were to be reported in aggregate format in the researcher's dissertation document, at professional conferences, and possibly in peer-reviewed journals in the future.

Participants submitted their responses to the online survey via the Qualtrics platform. An email thank you was submitted to state liaisons who were encouraged to share the researcher's appreciation for their participation in the survey.

State Participation and Response Rates. All six states involved in the study generated useable data for a total response rate of 16% which is acceptable for online survey methodology. Population total estimates and corresponding response rates are shown in Table 3-4.

Table 3-4

Survey Response Rates

Location	Population*	Number. of Responses	Response Rate	Percent of Responses
State1	300	20	6.6%	6.01%
State2	365	38	10.4%	11.41%
State3	180	73	40.5%	21.92%
State4	225	40	17.7%	12.01%
State5	425	63	14.8%	18.82%
State6	570	99	17.4%	29.73%
Missing Data	---	3	(0.0015%)	---
All	2065	336	16%	100%

Note: *Population estimates were provided by the liaisons in each participating state and represent the number of Extension professionals involved in Impact Evaluation who were invited to participate in the study.

Data from participants from all states included in this study—State1, State2, State3, State4, State5, and State6—are included in the "location" demographic factor. Given the acceptable rate of response, the data from this study were utilized to represent impact evaluation behaviors in the North Central region of Extension.

Data Cleaning and Item Non-response. Data from the survey were exported from Qualtrics to Excel, data review and cleaning were performed, and the data were uploaded into Stata to complete most of the statistical analysis. The data collected were stored in files on the

researcher's computer and backed up on flash drives and other storage devices for analysis. The researcher and methodologist were the only individuals who have access to the raw data.

A relatively low level of partial response and non-response suggested that study participants felt able to answer the survey questions. To illustrate, item nonresponse for the questions asking respondents about cultural and behavioral attributes rarely exceeded 1.0 – 2.5% of the responses, and often remained at or below 1%. The option to completely discard incomplete surveys was exercised during data cleaning when less than 20% of the survey was completed. Imputation was not used as a data cleaning or management strategy in this study. This approach was not taken because of the very low levels of nonresponse and missing data. Most questions with missing data had less than 1.5% of missing data.

It is important to note that the Stata statistical analysis software defaults to listwise deletion, which removes an individual's response from analysis if any of the items is missing a response. Pairwise deletion includes all answered items, even if one is missing. Pairwise is not available in all functions of Stata, therefore readers may notice that the *n* may vary in some analysis reporting. Thus, a few items with missing data showed an *n* of less than the 336 cases chosen as the official *n* for the study.

Data Analysis

Results from the survey were initially examined for descriptive statistics using Qualtrics. Data was then exported to Excel and cleaned to provide a file which could be imported, for analysis using Stata. Quantitative data were analyzed using Stata, Microsoft Excel, and the Qualtrics Research Suite software. Data were initially reviewed to determine the usability and quality of the data. Deletion of incomplete data sets that showed they were more than 20% incomplete was appropriate for this study especially due to the plan to use structural equation

modeling (SEM) for data analysis of the actual initial structural model which requires that data for all relevant parameters be included.

Data was reviewed using descriptive statistics and parametric inferential statistics. Measures of central tendency and tests of significance were calculated and examined; Cronbach's alpha was calculated to determine reliability; and variances (ANOVA) were calculated to explore relationships among factors that could not be included in the SEM.

Research Constructs and Factor Analyses

Analysis of data collected to measure constructs and variables included in the conceptual model was essential to calculate values to be used in the SEM analysis used to identify relationships and effects addressed by the study research question. Data analysis by construct or factor are described in the next section of this chapter.

Construct: Skill Levels of Competencies. Impact evaluation competencies rated for perceived skill level by study participants were analyzed to determine measures of central tendency and variance. The researcher computed an overall mean score for skill for competencies as well as for each competency. Items which showed the highest and lowest mean scores were identified and documented in data tables presented in chapter four.

Construct: Importance Levels of Competencies. Impact evaluation competencies rated for perceived importance by study participants were analyzed to determine measures of central tendency and variance. The researcher computed an overall mean score for importance of competencies as well as for each competency. Items which showed the highest and lowest mean scores were identified and documented in data tables presented in chapter four.

Construct: Competencies Importance vs. Reported Skill Level. A mean weighted discrepancy composite score (MWDS) was calculated based on the Borich model of needs

assessment (Borich, 1980; McKim, 2013). This score was used to test the SEM model to determine if it was an indicator of perceived behavior control. It was not significant when included in the model and, so, was omitted from the final structural model estimation. A MDWS may be calculated for each impact evaluation competency to identify future specific needs for training or staff development.

Construct: Culture Regarding Impact Evaluation. Items in the culture subscale were analyzed and a composite mean calculated to address the construct of impact evaluation culture. The composite score was utilized in the SEM path analyses. Descriptive statistics were also calculated for each culture item. Composite scores of three subcategories of the culture subscale were examined during SEM analysis to determine if the subcategories could be used as stand-alone factors in the model; they were not found to be significant. The three categories in this scale were: collaboration and problem solving; risk taking; and participatory decision making.

Constructs: Theory of Planned Behavior (TPB) Endogenous Factors. Composite means were calculated for each of the endogenous factors that were part of the core TPB section of the study conceptual model. These included: attitude (att), subjective norm (norm), perceived behavior control (pbc), behavioral intent (bint), and actual impact evaluation behaviors (actbx).

Construct: Number of Role(s) in Impact Evaluation. The importance of the role(s) performed as part of the impact evaluation process in Extension was included as a construct in the conceptual model based on literature review of related role and organizational role theories. Roles were combined into a composite score that represented the number of roles that each participant engaged in during impact evaluation. Role categories were not addressed individually for this study; completion of related research is one recommendation given in chapter five.

Construct: Teamwork. The construct of teamwork was included as a factor in the initial theoretical concept model based on literature review of evaluation capacity building, organizational culture and learning organization theory. Teamwork was measured on a sliding scale which directed participants to indicated what percentage of impact evaluation activities they performed as part of a team and what percentage was on-their-own. The data collected were on a continuous scale which was useable in the SEM analyses.

Individual and Organizational Demographic Factors. Composite means and relevant statistics for each demographic factor were calculated for inclusion in the SEM analyses. Categorical variables that could not be converted to continuous or dichotomous variables were not included in the SEM analyses. Descriptive analysis and regression analysis via ANOVA were performed to determine the levels of variation and correlation among those variables and to provide additional information to complement, contrast, and/or explain the results of the SEM analyses.

Re-specification of the Conceptual Model

In order to proceed with SEM data analysis, it was necessary to re-specify the initial theoretical model to create an actual initial model based on the data review and cleaning done for this study. Categorical variables that could not be converted to continuous or dichotomous variables were removed from the model as were items found to lack reliability. The item which measured support was found to lack reliability due to inconsistent participant responses within states that had separate program planning/evaluation support units; some participants reported their state did not have such a unit when it actually existed and, in states without a separate unit, some reported that their state had such a unit.

The re-specified conceptual model is shown in Figure 3-2. Factors deemed unusable in an SEM path analysis were removed from the model. This methodological issue necessitated the removal of these factors: program area, position type, training types, type of program planning model, and location (state). The categorical variables of program type, position type, and location were regressed on the dependent variable using ANOVA to assess relationships not included in the Actual Initial model. The factor of support was also removed from the model due to reliability concerns. The remaining factors are shown in the Actual Initial SEM model below:

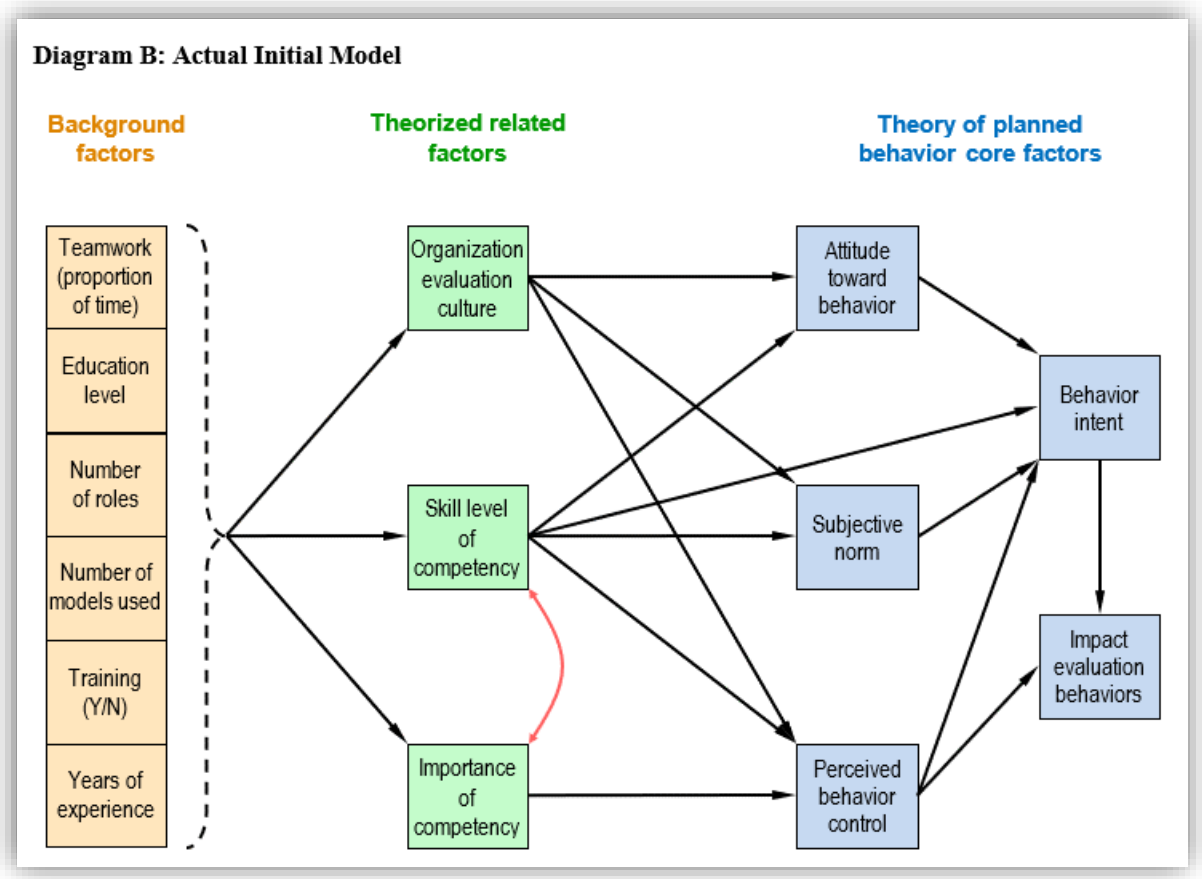

Figure 3-2. Initial model. Black arrows represent direct effects (pathways). The red double-headed arrow indicates correlated disturbance (error) terms.

Note that the model includes the factors deemed usable in SEM path analysis. Paths between factors indicate the theorized connections among the factors.

Research Question Analysis

The overall research question was: What are the relationships among the individual and organizational contextual factors that influence impact evaluation behaviors within Cooperative Extension? To address the question, SEM path analysis was performed on quantitative data collected that operationalized the factors identified in the conceptual model for the study.

SEM Analysis of Factors Influencing Impact Evaluation. Data from the competency skill and importance subscales, culture subscale, behavior subscales, and demographic factors measured were analyzed using structural equation model (SEM) path analysis to determine correlations and relationships among the factors. Prior to conducting the structural equation model (SEM) analysis, diagnostics were used to check for homodasticity and to determine that underlying assumptions of normality were met. Scatterplots were used for this portion of analysis. Cronbach's alpha was conducted to test for reliability and measures of internal consistency for each subscale; alpha values are reported earlier in this chapter in Table 3-1. For each major research construct—competencies, culture, and behavior—measures of central tendency and the amount of variance accounted for by each data section were calculated. Composite scores for competency skill, competency importance, impact evaluation culture, and Theory of Planned Behavior factors were utilized in SEM analyses. Demographic data were also used in the SEM data analysis. Only data in a format that could be converted to a continuous variable format were utilized in the SEM path model analyses. SEM analyses involved repeated iterations of testing for model fit using recommended processes for determining significance and strength of direct and indirect effects of factors in the various paths represented in the structural model. A measurement model was not created because sample size was too small (n = 336) to

allow for confirmatory factor analysis given the number of items (in the survey. Results of the SEM path analysis are shared in detail in chapter four of this document.

Qualitative Data Analysis

Patton (1990) described qualitative content analysis as a data reduction and sense-making process which examines data to identify core themes and to interpret meaning. Hsieh and Shannon (2005) noted that, if one records frequencies to determine recurrent themes, that "summative approach to qualitative content analysis goes beyond mere word counts to include latent content analysis. Latent content analysis refers to the process of interpretation" (p. 1283).

Qualitative data from the two open-ended questions were analyzed using coding strategies recommended by Saldaña (2013) and others (Layder, 1998; Boyzatis, 1998). Data were coded, recoded, sorted, annotated, re-sorted, and analyzed both by hand and using Excel and Word documents. Initial data review involved "pre-coding" (Layder, 1998) which included circling, highlighting, and color-coding participant quotes or phrases that were deemed worthy of special attention (Boyzatis, 1998). This pre-analysis review was done using printouts of exact transcripts of participant responses to the open-ended questions; pre-coding markups were then transferred to an Excel data file and saved as a revised, data file named "Qual-data--pre-code--1-2019" to organize and keep track of qualitative data handling during subsequent phases.

Qualitative Codebook. As initial coding evolved into structural coding, this researcher kept a record of the emergent codes in table format in a Word document. The number of codes accumulated relatively quickly during the initial stages of coding and changed as categories and subcategories emerged. The qualitative codebook table contained codes and subcodes, their abbreviations, content descriptions of each, a brief example of actual text from the data collected, and numbers assigned to each code and subcode.

Qualitative Coding and Analysis Strategies. Sub-coding was added early in the process of initial coding as a strategy to better organize the data into categories and related subcategories. The categories and subcategories tended to align with the constructs measured in the quantitative portions of the instrument in many cases; however, additional categories and subcategories emerged with subsequent rounds of recoding. Addition of subcategories was consistent with a structural coding approach. The categories which emerged were consistent with the conceptual and theoretical framework chosen a priori as a basis for the study's inquiries.

Second cycle coding is recommended by Saldana who noted that applying two complementary coding methods sequentially can yield a richer perspective of the same data set (p. 63). Second cycle coding employed for this research study involved pattern/focused coding to revisit the categorization of data done in the first cycle. Elemental, structural coding was undertaken as a second first-level approach to identify topics and summarize the statements made by participants in answer to the open-ended questions. Structural coding is most appropriate when analysis employs application of content-based or conceptual phrases to data obtained in response to a specific research question used to frame the inquiry topic (MacQueen, McLellan-Lemal, Barolow, & Milstein, 2008, p. 124). Structural coding is used to both code and initially categorize data in order to explore text and identify conceptual phrases' commonalities, differences, and relationships (Saldaña, 2013, p. 84).

Simultaneous coding was also utilized when needed; Saldaña noted that simultaneous coding is "appropriate when the data's content suggests multiple meanings that necessitate and justify more than one code," (p. 80). Limited use of simultaneous coding is recommended to avoid adding unnecessary categories and, thus, facilitate more succinct analysis of data.

Saldaña (2013) noted that some methodologists do not recommend combining qualitative data with data obtained via quantitative means; however, he posited that numeric representation of qualitative data is acceptable "when appropriate—as a supplemental heuristic to analysis" (p. 63). Thus, merging qualitative and quantitative data may illuminate and provide more in-depth explanation of quantitative results (Brannen & O'Connell, 2015). As well as providing a richer, more descriptive narrative analysis, qualitative data also offer a way to examine outliers that may be incongruent with the patterns found in the quantitative survey results.

Rather than discarding the more engaging exceptions to the core categories simply as outlier data with little to offer, those data were revisited to more thoroughly explore them narratively, as recommended by some narrative scholars (Frank, 2010; Riessman, 2008; and Sparkes & Smith, 2012). This exploration allowed for the potential identification of constructs or themes that had not been included in the original a priori conceptual model for the study. Divergent views, negative cases or outliers—however you choose to label them—provide a rich source for further analytic thinking.

Validity and Reliability of the Instrument

The instrument was created using previously validated scales and items where possible, underwent expert review to ensure construct validity, and clarity of instructions and wording of items. Factors regarding validity are discussed earlier in this chapter. Internal reliability was assessed using Cronbach's alpha following data collection.

Cronbach's Alpha coefficients for each subscale as well as the overall instrument were calculated and examined. Results of the alpha values for the endogenous factors to be included in the SEM analysis are reported in Table 3-5. Cronbach's alpha is a coefficient of reliability used to measure internal consistency and indicates how closely a set of items in a group are related. It

is considered to be a measure of scale reliability (DeVellis, 2016). Alpha values above .7 are considered acceptable, values above .8 are judged very good, and values above .9 are considered excellent. Values for items and the scale overall for this study were between .78 and .833, indicating reasonably high internal consistency.

Table 3-5

Cronbach's Alpha Values for Endogenous Factors to be Included in SEM Analysis

Factor/ Variable	Observations (n = x)	Average Interitem Covariance	Cronbach's Alpha (α)
cult (impact evaluation (IE) culture)	334	.1297446	0.8300
comp_s (competency by skill level)	335	.1333161	0.8070
comp_i (competency by importance)	335	.1460295	0.8244
att (attitude toward IE)	336	.1127807	0.7826
norm (subjective norms)	336	.1156451	0.7932
pbc (perceived behavioral control)	336	.1288827	0.8068
bint (behavioral intent)	336	.1060538	0.7825
actbx (actual IE behaviors)	336	.1302909	0.8328
Test scale average values (overall)		.1253464	0.8283

Reporting of Results

Research results are reported in chapter four of this document in a format to propose answers for the overall research question and to explore each of the research constructs/factors. Results are interpreted and implications discussed in chapter five through the lens of the concept model created for this research. Relationships of significance among the variables are examined and discussed. Qualitative results are included when appropriate to provide additional insight. Results are reported and reviewed in relationship to previous research, including recommendations for future research. The researcher includes suggestions for impact evaluation practice and for proposed evaluation capacity building strategies designed to strengthen impact evaluation competencies and evaluation culture within the Extension service.

Limitations

Potential limitations of the research include the fact that the list of impact evaluation competencies, although finalized by the American Evaluation Association in 2018, may change and may not reflect the exact competencies needed by a specific Extension state organization depending on their current needs and priorities.

The SEM modeling process assumes that an a priori model is properly specified. If relevant variables have been omitted, it is judged to have specification error. As is the case with missing data and non-normality of data, omitted variables have the potential to affect inferences from data analyses. To guard against this type of error SEM modeling was preformed using parceled variables to rule them out as additional factors, constructs and variables were validated using theory from previous research, and qualitative data were analyzed to determine if additional variables or factors could be identified as influencers on the target behavior of impact evaluation in this study.

Although organizational culture is measured using a previously validated scale, the factor of current organizational climate which involves the more temporal contextual factor regarding organizational support and/or lack of support for doing impact evaluation may also affect relationships in the model. Organizational climate was not measured and could potentially be a missing factor in the model. Climate may vary based on individuals' current workload, current physical and/or mental health status, current level of stress, and other temporal conditions.

Non-probability sampling, which was used for this research, can possibly result in a disproportionate image of the study's target population (Howell, 2006; Wells-Marshall, 2012). This sampling technique may limit generalizability of study results for the population.

Summary

Chapter three described the methods and procedures used in this study including the research design; SEM model specification and construct definitions; population and sampling plan; instrumentation; data collection strategy; and data analysis methods including structural equation modeling for quantitative data and thematic analysis for qualitative data to answer the research question.

A primarily quantitative, descriptive online survey was used for this study. Qualitative data were also collected, making the study a mixed method, non-experimental design. The data collection instrument was reviewed by experts in impact evaluation within Extension and higher education prior to the study. Data collection was accomplished according to accepted ethical standards of research with human subjects.

Congruent with the expanded Theory of Planned Behavior conceptual model for the study, the behavioral measures employed in this study included: attitude; perceived behavioral control; behavioral intentions; subjective norm; culture of evaluation; years of experience within Cooperative Extension; perception of competency importance and level of expertise; educational background; type of position within Extension; role(s) in impact evaluation; Extension program area and position type; support for impact evaluation; teamwork; program planning models employed; and training regarding evaluation, (Ajzen, 1991; Ajzen & Fishbein, 2005; Braverman & Engle, 2009; Conner & Armitage, 1998; Guion et al., 2007; Harder et al., 2010; Lambur, 2008; Lamm & Israel, 2013; McClure et al., 2012; Morford et al., 2006a; Workman & Scheer, 2012; Vengrin, 2015).

Research results are reported in chapter four of this document in a format to propose answers for the overall research question and to explore each of the research constructs/factors.

Results are interpreted and implications discussed in chapter five through the lens of the concept model created for this research.

CHAPTER FOUR. RESULTS

The purpose of the study was to employ a comprehensive, systems approach to study the interrelated individual and organizational factors that affect the practice of impact evaluation (IE) in Extension. Factors included were a) perceived skill levels and importance of impact evaluation competencies; b) impact evaluation culture of the organization; c) impact evaluation behaviors and their antecedents; and d) individual and organizational demographic factors such as educational level, program area, years of Extension experience, training, program planning models used, amount of teamwork, and more.

Research Question

The research question explored in this study was: What are the relationships among the individual and organizational contextual factors that influence impact evaluation behaviors within Cooperative Extension? Constructs explored in the study are shown in Figure 4-1.

Research Design

This research was a quantitative, non-experimental correlational study of the attitudes, evaluation competencies, evaluation culture, evaluation behaviors, and individual and organizational demographic factors associated with Extension educators' impact evaluation practices in the North Central Region of the Cooperative Extension service.

A major feature of this research design was the decision to use Structural Equation Modeling (SEM) statistical analysis technique. The use of SEM to organize and implement this study facilitated simultaneous exploration of many constructs theorized to be involved in impact evaluation behaviors in the context of Extension educational program evaluation.

Theoretical Framework and Initial SEM Model

The extended, modified TPB model created for this study is presented in Figure 4-1. The initial structural model shown includes the original variables of Ajzen's TPB model plus three

additional constructs: a) competencies-skill, b) competencies-importance, and c) IE culture. The model also includes both individual and organizational demographic variables. Factors included in the original TPB model are shown in blue; additional endogenous factors in the model include organizational impact evaluation culture and competencies by perceived skill level and by importance; these are shown in green. Exogenous factors represent both individual and organizational demographic influences on impact evaluation behaviors and are shown in salmon on the left end of the model. This model was adapted from the preliminary theoretical model.

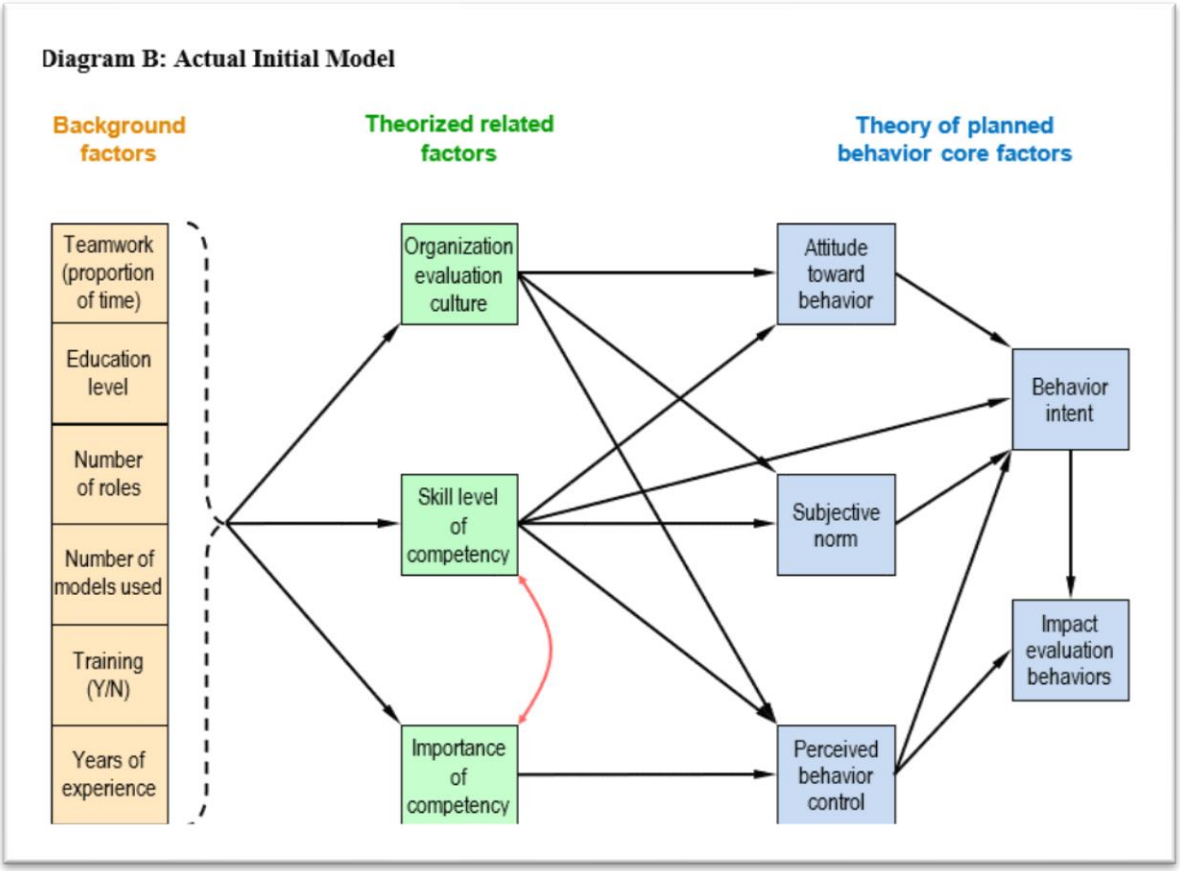

Figure 4-1. Initial model. Black arrows represent direct effects (pathways). The red double-headed arrow indicates correlated disturbance (error) terms.

Constructs and factors included were based on theory and research findings from academic literature. Theories incorporated, in addition to the TPB, included organizational culture theory, learning organization theory, role theory, competency and capacity building

literature, dissertation research and professional journal articles regarding evaluation in general and impact evaluation in Extension (Lamm, 2011; Ghimire & Martin, 2011; McClure, et al., 2012; Wells-Marshall, 2012; Workman, 2010; and Vengrin, 2016).

Impact evaluation planning, capacity building, and implementation are complex, interdependent, multi-faceted and dynamic and may best be understood using a comprehensive, systems approach (Urban, Burgemaster, Archibald, & Byrne, 2015; Trochim, Urban, Hargraves, Hebbard, Buckley, Archibald, & Burgemaster, 2012). Behavioral research has shown that an individual's attitudes, beliefs, and expectations are important predictors of behavior (Ajzen & Fishbein, 1980; Ajzen, 1991; Bandura, 1977).

One must also consider the effects of the context in which the individual is embedded regarding a specific target behavior such as impact evaluation. Context includes organizational culture and subjective norms; internal and external influences on individuals; and perceptions of organizational support for impact evaluation. Stame and Presti (2017) posited that evaluation strategies which focus on positive perspectives and success contribute to organizational learning. Collaboration and teamwork have been cited in evaluation research studies as essential to the effective design, implementation, and evaluation of educational intervention programs (Lamm, 2011; O'Sullivan, 2012; Patton, 2008; Roche, 1999). Research on the roles of evaluators in Extension is limited and has not included consideration of the number of roles and a relationship to actual evaluation behaviors. Research on number of roles performed and impact evaluation was not found in the literature; the factor was included based on role theory research in other contexts.

Descriptive Statistics

Parametric tests make assumptions about the parameters of the population distribution from which the sample is drawn. The assumption is often made that the population data are normally distributed. Data were found to be normally distributed for the all variables in the

study. Descriptive statistics were calculated for demographic factors regarding individuals and organizational characteristics. Results of those analyses are shown in Table 4-1 and Table 4-2.

Demographic Characteristics

Individual demographic characteristics of study participants are summarized in Table 4-2 on the following page. Organizational demographic characteristics reported by the participants are summarized in Table 4-3 on the page following individual characteristics.

Descriptive statistics for individual characteristics included the position respondents chose as their primary role in the organization. The largest percentage of respondents (60.9%) were from the category of county agents, program coordinators, and Extension educators; they are the frontline professional staff who present educational programs to clientele. The second largest percentage of professionals (27.2%) were from the specialist category which includes subject matter specialists at the state or regional levels. Individuals in this category lead the planning of educational programs and may serve as team leaders for planning impact evaluations. The third largest group of professionals (9.55%) were administrators at the state, department or district levels. The smallest groups (1.19%) indicated they were either program assistants or gave "other" as a response to the question. The four individuals who responded with "other" indicated they served in these types of positions: marketing and communications; program coordinator/specialist/administrator, and office professional (2).

Descriptive statistics for "program area" are also shown in Table 4-1 and indicate the category that respondents chose as their primary area of responsibility. The two largest groups of study participants were in the categories of Agriculture/Natural Resources (ANR) (32.74%) and Family and Consumer Science (35.12%). This is representative of the population for the study. The third largest group of participants had responsibilities for 4-H youth development (16.7%).

Table 4-1

Demographic Characteristics of Participants

Characteristic	***N***	**%**
Position currently held in Extension	335	100.0 %
County Agent / Extension Educator/Program Coordinator	204	60.9 %
Program Assistant (e.g. SnapEd, 4-H)	4	1.19%
Specialist (state or regional)	91	27.16%
Administrator (state, dept., district)	32	9.55%
Other	4	1.19%
Program Area	336	100.0 %
Agriculture/Natural Resources (ANR)	110	32.74%
4-H Youth Development	54	16.7 %
Family & Consumer Science (FCS)	118	35.12%
Community Resource Development (CRD)	30	8.93%
Other	24	7.14%
Years of Experience in Extension	334	100.0 %
< 5 years	56	16.77%
5 to 9 years	71	21.25%
10 to 4 years	68	20.35%
15 to 19 years	82	24.55%
20 to 24 years	19	5.69%
25 to 29 years	19	5.69%
30 to 34 years	10	2.99%
35 or more years	9	2.69%
Highest Degree Level	335	100.0 %
High School	1	0.3 %
Associate/2 yr	0	0.0 %
Bachelors	82	24.48%
Masters	188	56.12%
Doctorate	59	17.61%
Other	5	1.49%
Types of Training in Impact Evaluation*		100.0 %
On-the-Job experience	270	29.25%
Staff development, Face-to-face	247	26.76%
Staff development, Online	76	8.23%
Mentoring by administrator or colleague	120	13.00%
Part of Academic Course	46	4.98%
Grant-required	67	7.26%
Self-directed study	88	9.53%
Other	9	0.98%

Note. The sample for the study is N = 336; when there is missing data the observations reported are less than 336. There were minimal missing data in the useable responses for the study.
* For characteristics tagged with an asterisk (*) respondents were asked to check all that apply, making the "*n*" for this category the number of responses rather than the number of respondents.

Many Extension professionals involved in 4-H have dual responsibilities in another program area and 4-H. The fourth largest group of respondents indicated that they were responsible for Community Resource development (8.93%). The fifth largest group indicated "other" as their primary program area of responsibility (7.14%). Responses by those indicating "other" included the following, in order of most to least frequently mentioned: program and evaluation planning, Food and nutrition, administrative, horticulture, and geology and natural resources.

Organizational Demographics

Table 4-2

Demographic Characteristics of Organization as Perceived by Individuals

Characteristic	***N***	**%**
Program Planning Model(s) Used*	369	100 %
Logic models	282	76.4 %
Bennett's hierarchy	7	1.9 %
Kirkpatrick's levels of change	68	18.4 %
Other**	12	3.25%
Support for IE – Separate Unit	335	100 %
Yes	236	70 %
No	99	30 %
Teamwork – Proportion of IE Done in Teams	335	Mean %
Teamwork	187	56 %
On-your-own	148	44 %

Note: * Respondents were asked to check all that apply, making the "n" for this category the number of responses rather than the number of respondents.
** Not all respondents who answered "other" provided a description of their "other" response.

Years of experience. The years of experience in Extension ranged from 2 weeks (0.06) to 40 years. The distribution of experience is shown in the graph in Figure 4-2. The distribution over years of experience was a) 0 to 5 yrs. – 20.9%, b) 6 to 10 yrs. – 20.7%, c) 11 to 15 yrs. – 21,8%, d) 16 to 20 yrs. – 20.1%, e) 21 to 25 yrs. – 7.5%, f) 26 to 30 yrs. – 4,2%, g) 31 to 35 yrs. – .9%, and h) 36 to 40 yrs. – 1.8%. The distribution is relatively even between zero and 20 years

of experience and shows that over 80% of Extension professionals have between zero and 20 years of experience. A graphic representation of the distribution of years of experience for the study sample population is shown in Figure 4-2.

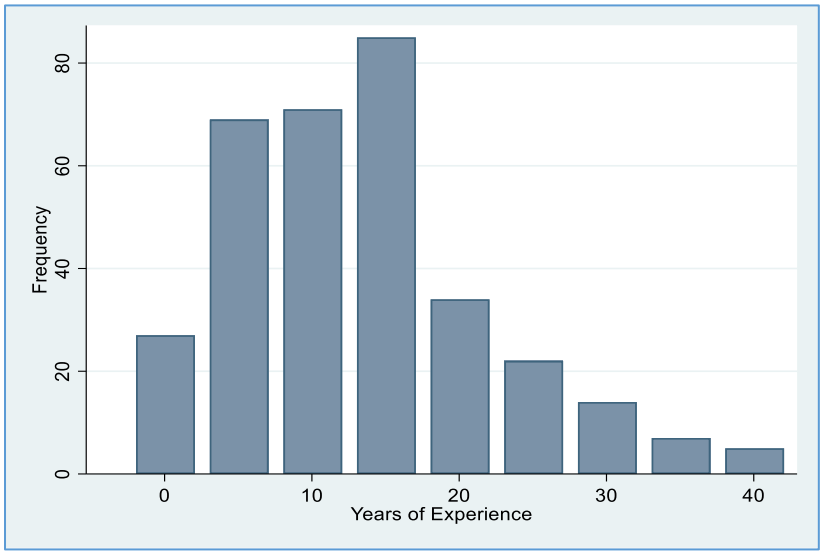

Figure 4-2. Years of experience in Extension.

SEM Constructs Analyses

Data collected to describe the constructs chosen for inclusion in the initial model for the study were analyzed to provide input for the SEM analyses and to examine each construct in detail. Constructs and factors and the measurement strategy for each were described in detail in chapter three. Data analyses for each construct is described in the next section of this chapter.

Data analysis: Impact evaluation competencies

Impact evaluation competencies were ranked by Extension professionals in two areas: 1) perceived skill level and 2) perceived level of importance. Results are presented in this section of chapter four. Competencies were subdivided into five domain areas, based on the current AEA evaluator competency list (AEA, 2018): 1) professional practice (pp), 2) methodology (mth), 3) context (ctx), 4) planning and management (pam), and 5) interpersonal (int).

In addition to examining perceived skill and importance levels, a mean weighted discrepancy score was calculated for each competency, according to the Borich method which is used to identify areas for staff development and training needs.

Competencies by Skill Level. Top-ranked competencies by perceived skill level are shown in Table 4-3.

Table 4-3

Highest Ranked Impact Evaluation Competencies by Perceived Skill Level

Competency (highest skill level first)	Mean	Standard Deviation	N	Rank
Act ethically when planning, doing, and reporting impact evaluation. (pp)	3.695	.479	335	1
Engage and listen to diverse perspectives (e.g. stakeholders) during impact evaluation. (int)	3.169	.605	334	2
Consider public value of the educational program being evaluated. (ctx)	3.167	.465	335	3
Collect data from accessible sources using appropriate, ethical procedures. (mth)	3.153	.588	334	4
Communicate impact evaluation information and results in timely, effective ways. (int)	3.123	.587	335	5

Note. Information in parentheses after the text of each competency refer to the domains specified by the American evaluation Association (2018): a) professional practice (pp), b) methodology (mth), c) context (ctx), d) planning and management (pam), and e) interpersonal (int).

The impact evaluation competencies ranked highest by perceived skill level included competencies from the following domains areas: professional practice (ranked 1), methodology (ranked 4th), context (ranked 3rd), and interpersonal (ranked 2nd and 5th). One could argue that the competency ranked 5th also relates to planning and management given that it includes the word "timely" as well as addressing effective communication.

In contrast, competencies rated lowest according to perceived skill levels are shown in Table 4-4. Note that the lowest skill level is listed first.

Table 4-4

Lowest Ranked Impact Evaluation Competencies by Perceived Skill Level

Competency (lowest skill level first)	Mean	Standard Deviation	N	Rank
Design impact evaluation studies to document evidence of change. (mth)	2.587	.691	334	21
Create and manage a feasible impact evaluation plan, budget, resources, and timeline. (pam)	2.594	.772	333	20
Analyze qualitative data appropriately. (mth)	2.721	.745	334	19
Understand the knowledge base of impact evaluation (theories, models, methods, and tools). (pp)	2.731	.598	335	18
Understand and use appropriate methods for impact evaluation including quantitative, qualitative, and mixed methods. (mth)	2.742	.726	334	17

Note. Information in parentheses after the text of each competency refer to the domains specified by the American evaluation Association (2018): a) professional practice (pp), b) methodology (mth), c) context (ctx), d) planning and management (pam), and e) interpersonal (int).

The lower ranked of the twenty-one (21) competencies include competencies in these domain areas: a) methodology – ranked last (21st), 19th, and 17th; b) planning and management – ranked 20th; and c) professional practice – ranked 18th. The majority of lowest perceived skills were from the methodology domain. This was consistent with the qualitative data themes found to be of most need to do effective impact evaluation, including "how to measure change" and "training in impact evaluation methods."

Competencies by Perceived Importance. Impact evaluation competencies by importance were ranked; the top five are shown in Table 4-5. Overall, competencies ranked by perceived importance had higher mean scores than those ranked for level of skill. This indicates there may be a gap in skill to perform important impact evaluation tasks.

Table 4-5

Highest Ranked Impact Evaluation Competencies by Perceived Importance

Competency (highest importance first)	Mean	Standard Deviation	N	Rank
Act ethically when planning, doing, and reporting impact evaluation. (pp)	3.907	.300	335	1
Communicate impact evaluation information and results in timely, effective ways. (int)	3.770	.435	335	2
Consider public value of the educational program being evaluated. (ctx)	3.748	.474	334	3
Use impact evaluation evidence/results to determine program effectiveness. (meth)	3.731	.476	335	4
Interpret impact evaluation findings relevant to the situation/context. (ctx)	3.724	.486	334	5

Note. Information in parentheses after the text of each competency refer to the domains specified by the American evaluation Association (2018): a) professional practice (pp), b) methodology (mth), c) context (ctx), d) planning and management (pam), and e) interpersonal (int).

The five competencies ranked highest by perceived importance represent the professional practice, methodology, context, and interpersonal domains in the scale. Two items were ranked in the top five for both importance and skill level--acting ethically when doing impact evaluation and the need to consider public value. Respondents judged these competencies to be of high importance and rated their skills in both quite highly. The competency regarding communication of impact evaluation "information and results in timely, effective ways," was also ranked high in importance. Three of the items, ranked 3rd, 4th, and 5th, dealt with competencies directly related to using impact evaluation results to improve or communicate the value of programs.

The lowest ranked competencies by importance are presented in Table 4-6. The lowest ranked items are listed first. Note that there was a tie in items ranked fifth from the bottom, in the 17th place overall, for the twenty-one items on the competency scale.

Table 4-6

Lowest Ranked Impact Evaluation Competencies by Perceived Importance

Competency (least important first)	Mean	Standard Deviation	N	Rank
Create and manage a feasible impact evaluation plan, budget, resources, and timeline. (pam)	3.141	.802	334	21
Design impact evaluation studies to document evidence of change. (meth)	3.302	.787	334	20
Understand the knowledge base of impact evaluation (theories, models, methods, and tools). (pp)	3.411	.561	335	19
Encourage constructive interaction and teamwork throughout the impact evaluation. (int)	3.419	.652	333	18
Understand and use appropriate methods for impact evaluation including quantitative, qualitative, and mixed methods. (mth)	3.434	.634	334	17*
Analyze quantitative data appropriately (mth)	3.434	.718	334	17*

Note. *Tied in the rankings. Note 2. Information in parentheses after the text of each competency refer to the domains specified by the American evaluation Association (2018): a) professional practice (pp), b) methodology (mth), c) context (ctx), d) planning and management (pam), and e) interpersonal (int).

The lowest ranked impact evaluation competency, by importance, was also ranked near the bottom by skill level. The competency, "Create and manage a feasible impact evaluation plan, budget, resources, and timeline," was clearly seen as less important by most Extension professionals who also ranked their skill level for the competency as low. The lower ranking of this competency by both skill and importance may indicate that professionals do not attend to developing the skill due to their perception of its importance. Or, could their perceived lack of skill in the competence affect how they ranked it in importance? Other lower ranked competencies that corresponded to low ranked perceived skill levels included understanding methodology appropriate for impact evaluation and data analysis competencies.

Data analysis: Culture of Organization regarding Impact Evaluation (IE)

The construct of impact organizational culture was measured using an adaptation of the culture subscale of the ROLE instrument by Preskill and Torres (1999). The highest ranked

statements used to assess impact evaluation culture of the organization are shown in Table 4-7. Survey respondents were asked to consider their organization's culture regarding impact evaluation, specifically, when answering the questions.

Table 4-7

Highest Ranked Impact Evaluation Organizational Culture Indicators

Impact Evaluation Culture Indicator (most important first)	Mean	Standard Deviation	N	Rank
My Extension colleagues and I respect each other's perspectives and opinions.	4.408	.716	333	1
Extension colleagues tend to work collaboratively with each other.	4.373	.776	334	2
My Extension colleagues and I operate from a spirit of cooperation, rather than competition.	4.324	.851	333	3
We ask each other for information about work issues and activities.	4.222	.702	334	4
We use data/information to inform our decision-making.	4.030	.894	335	5

Organizational impact evaluation culture items ranked highest emphasize teamwork and collaboration. This is supportive of the value of working together in the organization to complete impact evaluation.

Data-driven decision-making is also ranked highly and indicates that impact evaluation has potential to strengthen programming if results are made available for review and use. The perception that individuals are respected in for their perspectives and viewed positively as cooperators is a good indicator of a healthy, successful context for performing impact evaluation. Further, this may be an indication that espoused values of respect and teamwork are, indeed, practiced regarding working on impact evaluation.

Impact evaluation organizational culture indicators ranked the lowest are provided in Table 4-8. These indicators of culture show a picture of the organization that is not as positive as the highest ranked indicators. The incongruity of responses is puzzling. Further examination of the data is warranted to determine what other factors may influence whether respondents felt valued and respected.

Table 4-8

Lowest Ranked Impact Evaluation Organizational Culture Indicators

Impact Evaluation Organization Culture Indicator (least agreed with first)	Mean	Standard Deviation	N	Rank
In meetings we are encouraged to discuss the values and beliefs that underlie our opinions.	2.885	1.096	334	27
We are encouraged to offer dissenting opinions and alternative viewpoints.	2.958	1.106	334	26
We are confident that mistakes or failures will not affect us negatively.	3.212	1.098	333	25
Extension administrators make decisions after considering the input of those affected.	3.282	1.098	333	24
Extension administrators view individuals' capacity to learn as the organization's greatest resource.	3.298	1.056	332	23

Data reviewed in this section raised more questions and underscored the need to explore the data using SEM structural equation modeling to determine relationships among the factors and, specifically, to scrutinize the relationships between impact evaluation culture and others.

SEM Model Analysis

Variables. The actual initial SEM model adapted from the theoretical model specifies four types of constructs of interest for this study: Competencies, Organizational culture, Theory of Planned Behavior (TPB) factors (attitude, subjective norm, perceived behavior control, behavioral intent, and actual behaviors), and demographic variables at the individual and organizational levels. Those constructs modeled with latent variables were included in the model

as observed variables because they unable to be included in a measurement model as latent variables. Other constructs were modeled by single observed variables represented in the a priori SEM models.

Categorical Variables. The variables of location (state), program area, position type, were not included in the SEM path analysis because of their format (citation). These variables instead were analyzed using multiple regression (ANOVA); results for each ANOVA analysis are shown in later in this chapter.

Data Analysis Using Structural Equation Modeling (SEM)

Penke & Deary (2010) noted that "SEM is a powerful statistical tool to analyse complex relationships in multivariate datasets," (p, 1659).

Sample size recommendations for employing SEM include: 1) a minimum sample size of 100 and preferably > 200 and 2) a 10:1 ratio for cases to parameters estimated. This study produced a sample size of 336 (n=336) and had a sample case to parameter ratio of over 10:1 (336:19 or 17.6:1) for the original specified model. The adjusted structural model created through multiple iterations of testing model fit included twelve parameters that were significant, for a ratio of 336:12, equivalent to approximately 28:1 which is exceeds acceptable guidelines.

SEM Initial Model Respecification, Estimation and Testing. The preliminary theoretical conceptual model was re-specified once it was determined that a measurement model could not be calculated due to the ratio of items to factors in the data collected. There were 143 items in the survey and nineteen factors in the initially specified model. Thus, the model was too complex to calculate an accurate measurement model which would have required a sample size of at least 1430, a ratio of ten participants to each item (1430:143 or 10:1), and the minimum for calculating an SEM measurement model for this study. The sample size for the study was n =

336 and, although the sample size larger than required to use SEM to estimate a structural model, it was not sufficient to estimate a measurement model.

The re-specified actual initial model is depicted in Figure 4-3.

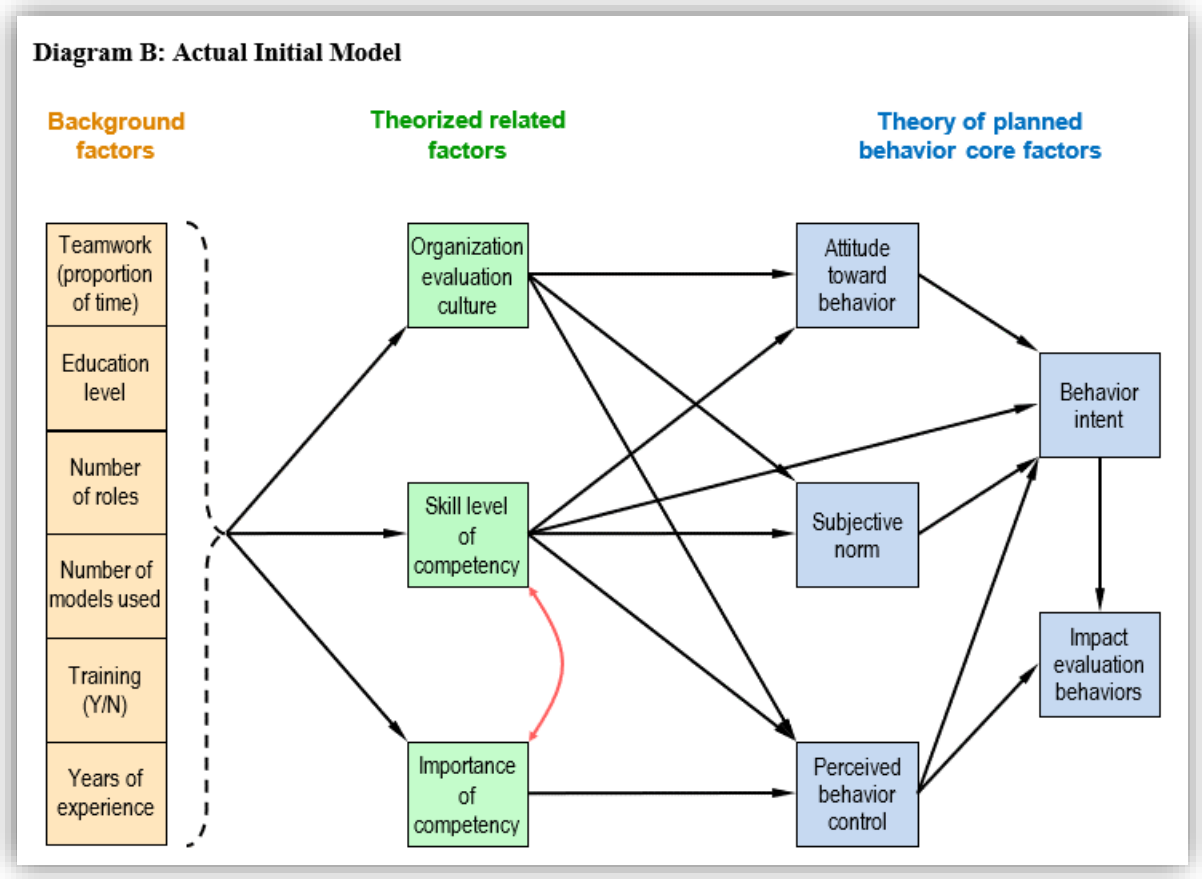

Figure 4-3. Initial model. Black arrows represent direct effects (pathways). The red double-headed arrow indicates correlated disturbance (error) terms.

The re-specified initial model omitted the following variables for the reasons stipulated: a) program area (*progarea*), position type (*position*), types of training (*ietrain*), and location/state (*location*) due to the fact they are categorical variables which could not be included in an SEM path analysis and b) *support* which was measured using a Y/N response to whether participants' states had a separate evaluation/program planning unit and which was found to lack internal reliability. The re-specified initial model had fourteen variables after the five noted above were

omitted. The model was tested in the initial SEM path analysis estimation. The estimation method used in all iterations of SEM analysis was that of maximum likelihood (ML) using the Stata software (version 15).

Actual Initial Model Analysis Results. Results of the initial model analysis showed that the factors of experience (*exper*) and IE program planning models used (*numb_iemodels*) were not significant in the model; they did not exhibit good local or global fit indices and no paths to the other factors to which they were connected in the original specified model were found. Thus, these factors were omitted from the model one at a time, following re-estimation which involved omitting paths first, and then omitting the factor when no paths to other variables could be found. Modification indices were consulted to determine if new paths were possible during the re-estimation process. Theoretical basis was also found to justify the omission of these two factors. Years of experience has not been found to be a factor when performing behaviors such as those associated with program impact evaluation when the task is new to most professionals which puts them on more equal footing regarding its performance. The number of program planning models used was not cited in the academic literature; the factor was initially included based on recommendations from expert Extension evaluators who reviewed the data collection instrument and was not justified in the model based on the estimation which did not find it significant.

Initial Model Fit Statistics. The results of the global fit statistical indices for the initial model are shown in Table 4-9. Guidelines for judging goodness of fit for the indices used are the following: non-significance for chi-squared; RMSEA $\leq$ .08 indicates good fit, neutral fit is between .08 and .10; SRMR $\leq$ .10 indicates good fit; CFI $\geq$ .90 indicates good fit; and TLI $\geq$ .90 indicates a good fit. It is important to note that with sample sizes over 200, the chi-squared test is nearly always found to be statistically significant (Kline, 2016). Thus, inflated value and significance make the chi-squared test a poor indicator of model fit given the large sample size.

Table 4-9

Global Fit Statistics for Initial Model

Test	Description	Value
χ^2	Chi-squared test	177.68*
CFI	Comparative fit index	0.883
RMSEA	Root mean squared error of approximation	0.102
SRMR	Standardized root mean square residual	0.074
TLI	Tucker-Lewis Index	0.778

* $df = 40$, $p < .001$

Fit statistics for the initial model show that the chi-squared test is significant which is due to the large sample size and, thus, does not indicate poor fit. The CFI is less than .90 and, so, indicates good fit; the RMSEA is .10 which indicates poor fit; the SRMR is .074 which indicates poor fit; and the TLI is 0.778 which indicates poor fit. These results, coupled with modification indices which recommended addition of new paths and deletion of others, were the basis for further estimation of alternate models to find one or more that better fit the data.

SEM Model Estimation Overview. The process of SEM model estimation and modification was done by making subsequent changes one at a time, as recommended by Boomsma and Hoogland (2001). Changing one parameter may have unpredictable and sometimes significant effects on a model, thus, a careful, iterative, single-step modification process aimed at identifying a model with better fit to our data was performed by adding or dropping parameters and/or paths from the model. Details of the analyses performed are provided in chapter four.

Iterations of SEM analysis of the model included many rounds of estimation in which path relationships found to be not significant were removed, relationships among variables not included in the original model were added based on modification indices, factors/variables found

to not be significant in the model were removed, and possible new factors based on the data collected were included to determine if they would improve the model. In all cases, only one change was made each time the model was re-estimated. New variables that were added included three subcategories of the culture construct from Preskill and Torres's ROLE culture scale, and five subcategories for both competency by skill level and competency by importance based on the American Evaluation Association domains for evaluator competencies. Parceling of the construct variables was done during model estimation to explore whether some factor/variable sub-categories could better explain relationships among the parameters chosen a priori for inclusion in the model. Parameter parceling also allowed the researcher to determine if sub-factors in parameters were significant. The testing of these "new" variables/factors did not result in statistical significance and no paths to other factors in the model were suggested in the modification indices. Because parceling yielded no significance for the parceled subconstructs; the potential new factors were not included in the model.

Table 4-10 shows the paths added to the model and those omitted from the model during the many iterations of re-estimation using Stata SEM calculations.

New paths that were added during the re-estimation of the model to explore which factors and paths were significant and congruent with the a priori theoretical/concept model included: a direct path from competency by skills level to actual IE behaviors, a direct path from competency by importance to subjective norm, and a direct path from competency by importance to attitude among the endogenous factors in the model.

Table 4-10

Model Modifications Made in the SEM Estimations of the Structural Model

Parameters added	Direct effects	norm ← comp_i att ← comp_i actbx ← comp_s
	Covariances between pairs of disturbance (error) terms:	cult, comp_i cult, comp_s att, norm att, pbc norm, pbc
Parameters removed	Direct effects	actbx ← pbc bint ← pbc cult ← edlevel cult ← num_roles cult ← num_iemodels cult ← train cult ← exper comp_s ← num_iemodels comp_s ← exper comp_i ← num_iemodels comp_i ← train comp_i ← exper

Literature supported the addition of each path that was added. For example, researchers have noted that, although one can have a strong intention to perform a behavior, without requisite skill the actual behavior may not be possible; this supports the addition of the direct path from competency by skill to the dependent variable of actual IE behaviors. Further, valuing competencies, measured by the perceived importance of IE competencies, was reported in literature as providing motivation which is related to the factor of attitude. Subjective norm includes shared values and expectations of what is deemed acceptable in a context; thus, if competencies are judged to be important, the factor of competency by perceived importance may affect subjective norm as indicated by the direct path to norm.

The most compelling change was the removal of two paths from the model which included direct paths from the endogenous variable of perceived behavior control (*pbc*) to

behavioral intent (intent) and to actual IE behaviors (*actbx*). This factor has been found to effect both behavioral intent and actual target behaviors in much research based on the Theory of Planned Behavior. However, some studies have found that the effects of perceived behavior control on intent and actual behavior were not significant, as in this study. This finding will be discussed in detail in chapter five.

Care was taken to make decisions on model modifications based on both the theory underlying the study design as well as on statistical, empirical indices. The SEM statistical fit indices of atheoretical modification to increase model fit was not pursued. No paths were fixed during the model modification process. All parameters included in the model were estimated using maximum likelihood (ML). No equivalent models with similar fit were identified during the model estimation process.

Further, no irregularities occurred during model estimation; Boomsma and Hoogland (2001) noted that irregularities of concern could include indications of empirical under-identification, convergence problems, and inadmissible solutions. The model chosen was more parsimonious than some that were considered during the process of model estimation. Congruity with the study theoretical perspective presented earlier served to ensure model identification.

The following variables were also tested in the process of model estimation and found not to be significant; thus, they were removed from the model: a) exp (yrs. of experience), b) frequency of working with impact evaluation support unit staff, and c) a mean weighted discrepancy score (MWDS) composite for the gap between competency importance and skills.

In this SEM path analysis, variables were treated as observed rather than as latent variables. The only latent variables in the path model are the disturbances which are not specified and may represent missing factors or indicate that a construct was not measured well, thus

generating an error disturbance.. All variables included in the model were treated as observed variables because an SEM measurement model was not calculable given the small sample size relative to the number of items used to measure the constructs.

The final structural model included twelve observed variables. Endogenous variables were: actbx (actual IE behaviors), bint (behavioral intention), att (attitudes toward IE), norm (subjective norms regarding IE), pbc (perceived behavior control regarding IE), cult (organizational IE culture), comp_s (skill level of IE competencies), and comp_i (perceived importance of IE competencies). Exogenous variables were train (IE training, y/n), edlevel (highest academic degree), num_roles (number of roles filled in IE behaviors), and propteam (proportion of IE behaviors performed using teamwork). Other variables included in the original conceptual/theoretical model were not found to be significant and were dropped from the model.

SEM path analysis identified direct, indirect, and total effect sizes among variables within the model. Effect sizes are reported after the figure which depicts the final structural model

Disturbance Variances in the Model

In a causal model such as the SEM path analysis used in this study, disturbance variances are indicated by curved lines with arrows on each end connecting two endogenous variables. The disturbances show that there are *unmeasured* factors which affect the endogenous variables connected by the disturbance curves (Pearl, 2012). Disturbance variances in SEM are not synonymous with residuals in multiple regression (Kline, 2016).

Disturbance variances in this study exist between the following variables: a) between attitude and subjective norm (β = .4); b) between attitude and perceived behavioral control (β = .27); c) between subjective norm and perceived behavioral control (β = .5); d) between culture

and competency by skill (β = .28); e) between culture and competency by importance (β = .19; and f) between competency by skill and competency by importance (β = .29).

The disturbance variances in the model may also indicate that there may be some areas of misfit due to the missing factors implied by the covariances among the model factors. Some of the disturbance may be explained by referring to the theory underlying the model. For instance, perceived behavior control was found to have no direct or indirect effects on other variables. However, other variables had direct and indirect effects on perceived behavior control. The Theory of Planned Behavior researchers added the construct of perceived behavior control to their model to help explain the value of self-efficacy, contextual factors, and perceived skill for performing a behavior. If respondents were self-efficacious and had high levels of internal locus of control the factor may not be as relevant to behavioral intent. Ajzen's theory also focused on control of behaviors that were generally performed as a choice by individuals. If Extension professionals know impact evaluation is required, they may not feel true choice is an option.

Kline (2016) noted that the addition of each disturbance correlation to a model "costs" one degree of freedom and, thus, makes the model more complex. It also improves the fit of a model. Kline recommends that, when there are substantive reasons for specifying disturbance correlations, it is best to analyze the model with the terms than without them (p. 138).

The final structural model is presented later in this chapter in Figure 4-4 and shows the effects among the factors. Effects depicted are indirect or direct; total effect on the dependent variable is calculable and shown in Table 4-11 later in this chapter. Effects among the variables are explained after the model graphic and initial description of the most influential effects are briefly discussed. Further discussion of the effects in the model is found in chapter five.

Global Model Fit

Model fit was also assessed using Stata to calculate the global fit statistics of Chi-squared, RMSEA, TLI, SRMR and CFI. The results of the tests of global model fit for both the initial and the final models are reported in Table 4 –11. Obsolete global model fit statistics, including the NFI, IFI, and GFI, were not used and are, in fact, not included in the current version of the Stata software.

Table 4-11

Global Fit Statistics of Final Structural Model Compared to Initial Model

Test	Description of Test	Value	
		Final Model	Initial Model
χ^2	Chi-squared test	123.59*	177.683**
CFI	Comparative fit index	0.922	0.883
RMSEA	Root mean squared error of approximation (90% confidence interval)	0.093	0.102
SRMR	Standardized root mean square residual	0.070	0.074
TLI	Tucker-Lewis Index	0.854	0.778

Note. *(*df* = 32, *p* < .001); ** *(df*=40, *p* < .001)

Values for the final model shown in the global fit statistics Table 4-11 were interpreted using guidelines for each test. Note that fit statistics for the initial model are provided for comparison; the fit tests for that model were found to exhibit poor fit in most cases (see table 4-9). Most fit tests reported for the final model showed goodness of fit (CFI = .92), (SRMR =.07), and (TLI = .85). There was a significant chi-squared statistic, $\chi^2(32) = 123.59$, $p < .01$. It is prudent to note that, with sample sizes over 200 as in this study ($N = 336$), the chi-squared test is nearly always found to be statistically significant (Kline, 2016). Thus, inflated value and significance make the chi-squared test a poor indicator of model fit given the large sample size. Further, the RMSEA = 0.9 is higher than recommended; a RMSEA value within the range of

0.08–0.10 is seen as indicative of a fit which is neither good nor bad (Cangur & Ilker, 2015). Good fit does not ensure that model is correct, only that it is plausible. Most fit indices examined for the study model with best global fit chosen through iterative SEM analyses showed reasonable to *good* fit except for the RMSEA, which was inconclusive.

It is prudent to note that Kenny and McCoach (2003) showed that RMSEA improves as more variables are added to a model. Thus, had the model been able to accommodate all variables originally theorized in the specified model, including the categorical variables analyzed separately using ANOVA and found to have statistical significance regarding the dependent variable, the model may have exhibited a RMSEA value showing better model fit.

Local Model Fit

Tests of local model fit focus on the value of individual parameters. Two main tests were used to determine local model fit and to adjust the model to exhibit better fit. The main test consulted was that of parameter significance for z-tests examined to determine if the variable relationships were significant for $z > 2.0$ at the .05 level of statistical significance. Parameters with statistical significance were judged to be of good local fit and kept in the model. The z-test is more powerful and accurate than a t-test because the z-test uses the study population standard error whereas a t-test uses estimated standard error. Modification indices (MI) were also used to determine if making changes in a model might result in better fit. A new parameter may be a good candidate to add to a model if its MI is significant at $MI \geq 4$. Recommended practice is to consider only parameters with an $MI \geq 10$. The researcher worked with a SEM data analysis expert to go through the process of testing local and global model fit. Deleting statistically nonsignificant parameters increased degrees of freedom and resulted in a more parsimonious, simpler model.

Equation Level Goodness of Fit for the Final Model. Table 4-12 shows the equation-level goodness of fit calculations for the endogenous variables in the final model expressed as Cohen's R^2 effect sizes. Cohen's levels of effect sizes are interpreted as: .10, small/weak; .30, medium/moderate; and .50, large/strong (Kline, 2016). Thus, one can see that the final model exhibits the highest goodness of fit for the endogenous variable of behavioral intention (bint) with an R^2 large effect size of .299. This is followed by a medium effect size of .336 for the factor of attitude. The next highest levels of goodness of fit were actual impact evaluation behaviors (actbx) with a medium effect size of .286, subjective norm (norm) with a lower medium effect size of .241, and competencies by perceived skill (comp_s) with a low/medium effect size of .232. Perceived behavior control (pbc) also exhibited a low/medium effect size of .231. The lowest effect sizes in the model were competencies by perceived importance (comp_i) at R^2 = .133 and IE culture of the organization with a small effect size of R^2 = .027.

Table 4-12

Equation-level Goodness of Fit for the Final Model – Cohen's R^2

Endogenous variable	R^2
actbx	.286
bint	.599
pbc	.231
att	.336
norm	.241
comp_s	.232
comp_i	.133
cult	.027

The researcher notes that good fit indices show that variances in the variance-covariance matrix is well-represented by the model. This does not, however, imply any explanation of the amount of variance in the endogenous variables (Boomsma & Hoogland, 2001).

Parameter Estimates from the Model Solution for Structural Model. Standardized and unstandardized parameter estimates are shown in Table 4-8. All estimates were found to be significant at the $p = .01$ nominal level for the parameters chosen to be included in the model. Estimates are given for variances of the endogenous and exogenous variables. The estimates for these parameters for the final structural model estimated are given in Table 4-10.

In SEM regression-based analysis, different units and different scales are often used. For example, in this study one variable—experience in Extension--was measured using years and another—proportion of teamwork--used percentages. When measurement scales of independent variables are different, researchers must use standardized coefficients to interpret and compare their effects on the dependent variable.

Standardized coefficients (β) allow a researcher to interpret and compare the relative importance, or effect levels, of independent variables on the dependent variable in the SEM regression model (Glen, 2019). Unstandardized coefficients (*b*) are 'raw' coefficients produced by SEM regression analysis performed on original, unstandardized variable data. To determine the most interesting effects, a researcher must standardize the coefficients first, which means they are both in terms of standard deviations and, so, may easily be compared with each other. These calculations were done using the Stata statistical analysis program to produce the standardized and unstandardized coefficients shown in table 4-13.

Table 4-13

Parameter Estimates for the Final Structural Model – Direct Effects

Parameter	Standardized estimates (β)	Unstandardized estimates (b)
Direct effect to actual behavior from behavioral intent	.178	.176
Direct effect to actual behavior from compentency-skill level	.432	.719
Direct effect to behavioral intent from attitude	.478	.568
Direct effect to behavioral intent from subjective norm	.339	.381
Direct effect to behavioral intent from competency-skill level	.093	.157
Direct effect to competency-skill from proportion of teamwork	.127	.002
Direct effect to competency-skill from eduction level	.229	.133
Direct effect to competency-skill from number of roles	.278	.075
Direct effect to competency-skill from training (y/n)	.148	.133
Direct effect to attitude from competency-skill level	.309	.442
Direct effect to attitude from IE culture	.270	.233
Direct effect to attitude from competency-importance	. 216	.364
Direct effect to subjective norm from competency-skill	.194	.292
Direct effect to subjective norm from culture	.312	.283
Direct effect to subjective norm from competency-importance	.168	.299
Direct effect to culture from proportion of teamwork	.166	.004
Direct effect to competency-importance from propteam	.109	.001
Direct effect to competency-importance from education level	.145	.071
Direct effect to competency-importance from number of roles	.281	.064
Direct effect to perceived behavior control from comp-skill level	.353	.432
Direct effect to perceived behavior control from culture	.143	.105
Direct effect to perceived behav control from comp-importance	.128	.185

Note. Direct effects only are included in this table. Additional effects for this model include indirect effects, total effects, and covariance disturbance effects which are reported elsewhere.

Neither standardized nor unstandardized SEM regression coefficients are inherently better. The decision of which to choose depends on what is to be compared and the goals of the research. In this study, comparison of the independent variables and their effects on the

dependent variable of actual impact evaluation behaviors is the goal; thus, standardized SEM regression coefficients were used to determine effect sizes.

Covariance Matrix. The covariance matrix for the structural SEM model found to exhibit goodness of fit is provided in Table 4-11. Covariance is the measure of how two variables change and are, thus, correlated, with respect to each other.

The covariance matrix of the specified SEM structural model shows a comparison of data in two variable data sets. A covariance matrix, like many statistical matrices, is symmetric. The off-diagonal elements show the covariances of each pair of variables. The diagonal elements contain the variances of each variable; these variances show the distribution of the data around the mean. In SEM analysis structural models are evaluated by comparing two covariance matrices; the model with the best fit is calculated and shown in the software output.

Values from the covariance matrix are used by the Stata program to calculate the SEM structural model. Data were collected based on the theoretical foundation of the specified model. Thus, covariance-based structural equation modeling facilitates the testing of that theory. In the case of this study, the model was re-specified because a measurement model was not possible due to sample size being too small for that analysis. However, SEM path analysis was employed using the variables that could be validated and included in the re-specified model. The covariances of those variables were included in the model estimation and modification that resulted in the final structural model found to be of good fit. The covariance table shared in Table 4-14 represents all variables included in the final structural model.

Table 4-14

Covariance Matrix for Correlations Among the Measured SEM Variables

	comp_s	comp_i	cult	att	norm	pbc	bint	actbx	propteam	edlevel	num_roles	train
comp_s	.168758											
comp_i	.05839	.119355										
cult	.08067	.046978	.437531									
att	.114591	.080168	.159147	.341974								
norm	.089782	.066169	.16739	.205719	.380786							
pbc	.092162	.052245	.091698	.137908	.186748	.250847						
bint	.125735	.096335	.162402	.290477	.275899	.160062	.481548					
actbx	.143493	.04201	.079073	.151946	.092815	.079578	.175047	.463411				
propteam	1.56385	.043336	3.23583	3.03907	3.41471	2.80692	1.05569	-.08426	835.489			
edlevel	.10337	.060102	.025869	.108144	.061967	.06332	.132471	.146608	.738816	.499363		
num_roles	.226079	.16845	.017054	.274323	.280269	.20876	.436642	.36118	-4.53537	.340644	2.30316	
train	.047391	.012888	.02548	.045201	.087778	.06723	.049924	.067715	2.3264	.052069	.088423	.207431

Note. Variables represented in this table are: comp_s (competencies by self-rated skill level), comp_i (competencies by perceived level of importance), cult (organizational IE culture), att (attitude toward IE behaviors), norm (subjective norms regarding IE), pbc (perceived behavior control regarding IE behaviors), bint (behavioral intent to perform IE), actbx (actual IE behaviors), propteam (proportion of IE performed using teamwork),edlevel (highest degree attained), num_roles (number of roles filled when doing IE), train (training in IE – y/n).

Direct and Indirect Effects in the Model

Assessing the strength of direct and indirect relationships in the model can give insight into which factors might explain variation in the dependent variable and could be targeted for intervention through staff development or training to influence the relationship in a positive manner. One example of such a relationship is the causal effect (CE) between attitude (att) and behavioral intention (bint). To calculate the relationship, we add the DE + IE = .45 + 0 = .45 = CE. To calculate the total relationship (TR) we add DE + NA (non-causal associations) = .45 + (.4 *.34) = .45 + .136 = .586 = côrr(att, bint). Thus, the total relationship is a covariance/ correlation of .586; however, while it is important to account for noncausal associations, the causal relationships (direct, indirect, and total causal effects) are typically the primary focus for interpretation and reporting (Kline, 2016). Thus, the regression weight of .45, the direct causal effect, is the value most often interpreted and reported.

Other causal direct and indirect effects found in the regression weights for the model are reported in the graphic representation of the Final Structural Model on the next page. Effects are outlined and briefly assessed. These findings will be summarized and discussed in chapter five.

The most influential factors which may be calculated from the direct and indirect effects shown in the Final Structural Model are presented in Table 4-12 later in this chapter. The top seven total effects found were: 1) competency by perceived skill (β = .487), 2) behavioral intention (β = .179), 3) number of roles (β = .143), 4) education level (β = .116), 5) attitude (β = .085), 6) training (β = .072), and 7) proportion of teamwork (β = .072).

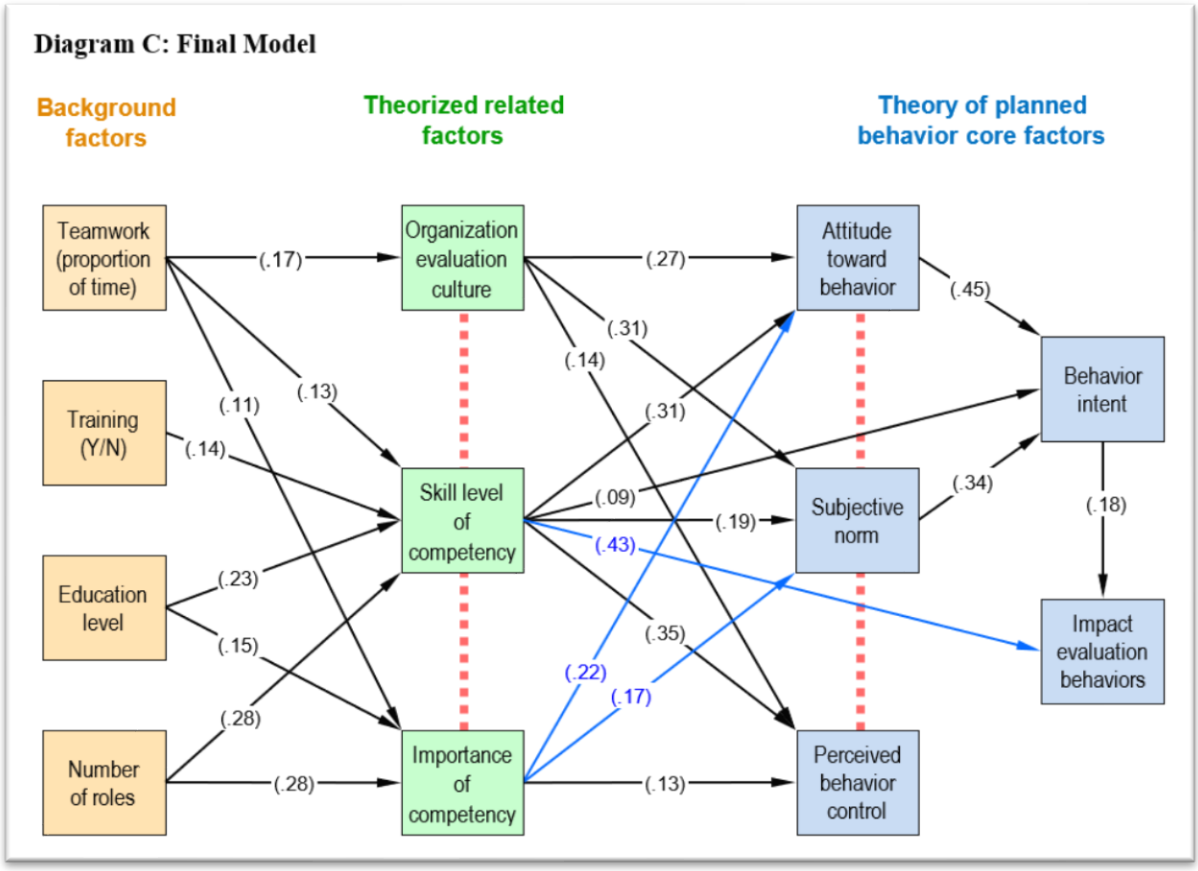

Figure 4-4. Final model after modification. Black arrows represent direct effects (pathways). The blue paths were added to the model to achieve acceptable fit. The standardized path coefficient estimates are shown in parentheses. The red dashed lines are used to indicate subsets of endogenous variables that have correlated disturbance (error) terms.

Factor Relationships and Effect Sizes

Standardized regression weights, or Beta coefficients (β), which are shown for each path in the model can be used to determine the effect sizes and relative importance of variables that had direct effects on an endogenous variable within the structural model. General guidelines for interpreting the relative size of standardized regressions weights are: .10, small/weak; .30, medium/moderate; and .50, large/strong (Kline, 2016). The relative effects may also be expressed as a percentage (Kline, 2016). In this study all Beta coefficients in the model were positive.

In this study, the variables directly influencing the dependent variable of actual impact evaluation (IE) behaviors were competency skill level with a medium level effect (β =.43) and

behavioral intent with a lower level effect (β =.18). The effect of competency skill level exceeded that of behavioral intent by about 139% (.43/.18= 2.39). Conversely, the effect of behavioral intent on actual IE behaviors is about 42% of competency by skill.

The variables of attitude (β =.45) and subjective norm (β =.34) show medium level effects on the variable of behavioral intent. The effect of attitude on behavioral intent exceeds that of subjective norm by approximately 32% (.45/.34=1.32).

Perceived competency skill level had a large effect (β = .43) on actual IE behaviors. Subjective norm showed a medium direct effect (β = .34) on behavioral intent and attitude showed a medium to large effect (β = .45) on behavioral intent. Training exhibited a small effect (β =.14) on only one variable, that of competency by skill. Thus, it was not one of the more influential variables in the model. Proportion of teamwork had small direct effects on both culture ((β =.17) and on competency by importance (β =.11), indicating that it has minor influence in the model; both paths were significant at $p \leq .001$.

To examine the effects in the model in more detail, post SEM analysis was performed using the Stata statistical software program to calculate total and indirect effects in addition to the direct effects shown in the final structural model (Figure 4-4). The results of the calculation of total effects of all variables represented in the model are shown in table 4-15. Calculations of all effects in the model—direct, indirect, and total--are given in table 4-16 to show comparisons within the model. Review of the various effect sizes is provided in narrative form following the two tables just mentioned. All effect sizes are judged assessed using the same guidelines as noted earlier with effect sizes and importance: .10, small/weak; .30, medium/moderate; and .50, large/strong.

Table 4-15

Total Effects of Final Structural Model Variables on Actual IE Behavior

Variable/Factor	Total Effect
Behavioral Intention	.17882
Attitude	.08547
Subjective Norm	.06059
Perceived Behavioral Control*	-----
Culture (Impact Evaluation)	.04199
Competency – Skill	.48700
Competency – Importance	.02859
Training	.07207
Education Level	.11603
Number of Roles	.14340
Proportion of Teamwork	.07184

Note. * There were no paths from perceived behavioral control to other variables; thus, no indirect, direct or total effects exist for that variable in the model.

Table 4-15, above, is show the total effects of all variables/factors in the model on the dependent variable (DV) of actual impact evaluation behaviors. Factors with the largest total effects were competency by skills level, behavioral intention, number of roles, and education level. Slightly smaller total effects were found for the factors of attitude, training, and proportion of teamwork. Factors with the smallest total effects in the model were subjective norm, culture, and competency by importance. Perceived behavior control had no total effect on the DV.

Total effects are significant at $p \leq .001$ for behavioral intention, competency by skill, attitude, subjective norm, education level, and number of roles. Total effects are significant at $p < .05$ for culture ($p = .002$) and training ($p = .002$). Total effects on the dependent variable overall were not significant for either competency by importance ($p = .006$) or proportion of teamwork ($p = .006$); however, the variables were significant in the model and showed significant direct and indirect effects on other model variables. No indirect or direct paths between perceived behavior control and actual behavior; thus, there was no significant effect reported for perceived behavior control.

Table 4-16

Effects of Variables/Factors on the Dependent Variable of Actual IE Behavior

Variable/Factor	Path	Direct Effects	Indirect Effects	Total Effect
Behavioral Intention	Direct path to actual behavior	.17882		.17882
Attitude	through behavioral intention	.47800		.08547
Subjective Norm	through behavioral intention	.33885		.06059
Perceived Behavioral Control*	No paths to actual behavior	-----	-----	-----
Culture	through behavioral intention		.23486	.04199
	through attitude	.27033		
	through subjective norm	.31176		
	to perceived behavioral control	.14268		
Competency – Skill	Direct path to actual behavior	.43219	.05480	.48700
	through behavioral intention	.09263	.21384	
	through attitude	.30970		
	through subjective norm	.19418		
	through perceived behavior control	.35274		
Competency - Importance	through behavioral intent		.15991	.02859
	through attitude	.21552		
	through subjective norm	.16791		
	to perceived behavior control	.12773		
Training	through behavioral intent		.04535	.07207
	through attitude		.04563	
	through subjective norm		.02874	
	through competency by skill level	.14798		
	to perceived behavioral control		.05220	
Education level	through competency by skill level	.22972		.11603
	through competency by importance	.14513		
Number or Roles	through behavioral intent		.13018	.14340
	through attitude		.14673	
	through subjective norm		.06897	
	through competency by skill level	.27793		
	through competency by importance	.28141		
	to perceived behavior control		.13398	
Proportion of Teamwork	to actual IE behavior		.07184	.07184
	through behavioral intent		.09511	
	through attitude		.10743	
	through subjective norm		.09446	
	through culture	.16550		
	through competency by skill level	.12688		
	through competency by importance	.10854		
	to perceived behavioral control		.08223	

Note. Endogenous variables are shown in blue and green; exogenous variables are in orange and yellow.
*The were no paths from perceived behavioral control to other variables; thus, it had no effects to report. However, the factor was significant in the model and other factors showed effects on it.

Additional Noteworthy Effects within the Model

Direct effects, indirect effects, and total effects of the model variables/factors on the dependent variable of actual IE behaviors and on other endogenous variables in the model are

shown in table 4-16 and compared in the next section of this chapter. All effects described were significant at $p \leq .05$ or better.

Behavioral intent shows only a small direct effect ($\beta = .18$) on one variable—that of actual impact evaluation behaviors. The direct effect in this case is equivalent to the total effect on the dependent variable. Three variables showed direct effects on the variable of behavioral intent—attitude had a medium/large effect ($\beta = .18$), subjective norm had a medium effect ($\beta = .34$), and competency by skill level had a small effect ($\beta = .09$).

The factor of competency by skill level showed the highest total effect ($\beta = .49$) and the most direct effects to other factors in the model, making it the primary influence over the dependent variable. Competency by skill level had direct effects on actual impact evaluation behaviors (medium/large, $\beta = .43$), on behavioral intent (small, $\beta = .09$), on attitude (medium, $\beta = .31$), on subjective norm (small, $\beta = .19$), and on perceived behavioral control (medium, $\beta = .35$). It was clearly the most influential factor overall in the model.

Behavioral intent was the second most influential factor in the model with one direct effect direct ($\beta = .18$) on actual impact evaluation behaviors. This effect was small, however, the overall total effect on the DV was the second highest in the model, indicating that the cumulative effects of other factors on behavioral intent was substantial.

The third highest total effect in the model was shown by the exogenous factor of number of roles, which had a small effect ($\beta = .14$) on the DV of actual IE behaviors. Number of roles had equivalent small/medium direct effects on competency by skill level ($\beta = .28$) and on competency by importance ($\beta = .28$). Number of roles had small indirect effects on three other endogenous variables within the model; an effect on behavioral intent ($\beta = .13$), on attitude ($\beta = .15$), and on subjective norm ($\beta = .07$).

The fourth highest total effect and level of influence in the model was the factor of education level, which showed two direct effects in the model—a small/medium effect on competency by skill level (β = .23).and a small effect on competency by importance (β = .15).

The factor with the fifth most amount of influence in the model was attitude toward the DV of actual impact evaluation behaviors , with a small total effect (β = .085). The effect of attitude on the factor of behavioral intent was significant with a medium/ large effect (β = .45).

The sixth highest level of influence in the model overall was that of training which had a small direct effect (β = .14) on only one factor—that of competency by skill level. The variable of training measured whether participants had training specifically in impact evaluation. Training also exhibited small indirect effects on four other factors in the model: behavioral intent (β = .045), attitude (β = .045), subjective norm (β = .028), and perceived behavioral control (β = .05).

Proportion of teamwork had a small total effect (β = .028), within the model although it exhibited direct effects on three other factors and indirect effects on five other factors. Direct effects were small on culture (β = .17), competency by skill level (β = .13), and on competency by importance (β = .11). Indirect effects were small on the factors of actual impact evaluation behaviors ((β = .07), on behavioral intent (β = .095), on attitude (β = .11), on subjective norm (β = .094), and on perceived behavioral control (β = .08).

Subjective norm showed a small total effect (β = .06) on the DV in the model and had only one medium direct effect to behavioral intent (β = .34). The factor had no other effects.

The factors of culture and competency by importance exhibited the lowest total effects in the model; both were small with culture at β = .04 and competency by importance at β = .028. Culture had small effects on four other factors; a direct effect on behavioral intent (β = .23) and indirect effects on attitude (β = .27), subjective norm (β = .31), and perceived behavioral control

(β = .14). Competency by importance had one small direct effect on behavioral intent (β = .16) and three small direct effects on attitude (β = .22), subjective norm (β = .17), and perceived behavioral control (β = .13).

Discussion of these and other effects reported in this chapter will be reviewed, interpreted, and discuss with respect to previous research, implication for future research and practice in chapter five.

Supplemental Analyses Using One-way ANOVA

The following analyses were done to provide information to help enrich and interpret the data analyses completed using SEM path analysis to better answer the research question.

Categorical exogenous variables that could not be included in the SEM analysis were analyzed using one-way ANOVA to determine if any significant relationships might exist among the factors and the dependent variable of actual impact evaluation behaviors (actbx). One-way analysis of variance (ANOVA) is used to compare the differences between two or more groups or categories of an independent variable with one dependent variable.

Separate one-way ANOVA tests were run to compare each of the following variables to the dependent variable of actual impact evaluation (IE) behaviors (actbx): a) program area (progarea) , b) position type (position), c) location/state (location), and d) separate evaluation department (sepevdept).

A one-way ANOVA was conducted to determine if the effect of program area on actual impact evaluation was different for different program area categories. There was a statistically significant difference between groups reported $F(4,334) = 10.66$, $p \leq .001$. Study participants were classified into five program areas: a) Agriculture and Natural Resources (ANR, $n = 110$), b)

4-H Youth Development (4-H, $n = 54$), c) Family and Consumer Science (FCS, $n = 118$), d) Community Resource Development (CRD, $n = 30$), and e) Other ($n = 24$).

A Tukey post hoc test indicated that actual impact evaluation (IE) behaviors were statistically significantly higher in the community resource development (CRD) category compared to the agriculture and natural resources (ANR) category (.66 ± .13 actual IE behaviors, $p \leq .001$). Actual IE behaviors were statistically significantly higher in the CRD category compared to the 4-H youth development category (.88 ± .15 actual IE behaviors, $p \leq .001$). Actual IE behaviors were also found to be significantly higher in the CRD category than in the Family and Consumer Science (FCS) category (.64 ± .13 actual IE behaviors, $p \leq .001$). Actual IE behaviors were also found to be significantly higher in the other category compared to the 4-H category (.61 ± .16 actual IE behaviors, $p \leq .05$). However, there were no statistically significant differences between the 4-H and ANR categories (-.22 ± .11 actual IE behaviors, $p = .234$), between the FCS and ANR categories (.01 ± .09 actual IE behaviors, $p = 1.0$), between the FCS and 4-H categories (.24 ± .11 actual IE behaviors, $p = .174$), between other and ANR categories (.38 ± .15 actual IE behaviors, $p = .081$), between the other and FCS categories (.37 ± .15 actual IE behaviors, $p = .098$), or between the other and CRD categories (-.28 ± .18 actual IE behaviors, $p = .545$). Figure 4-5 is a graphic representation of the adjusted predictions of actual IE behaviors by program area.

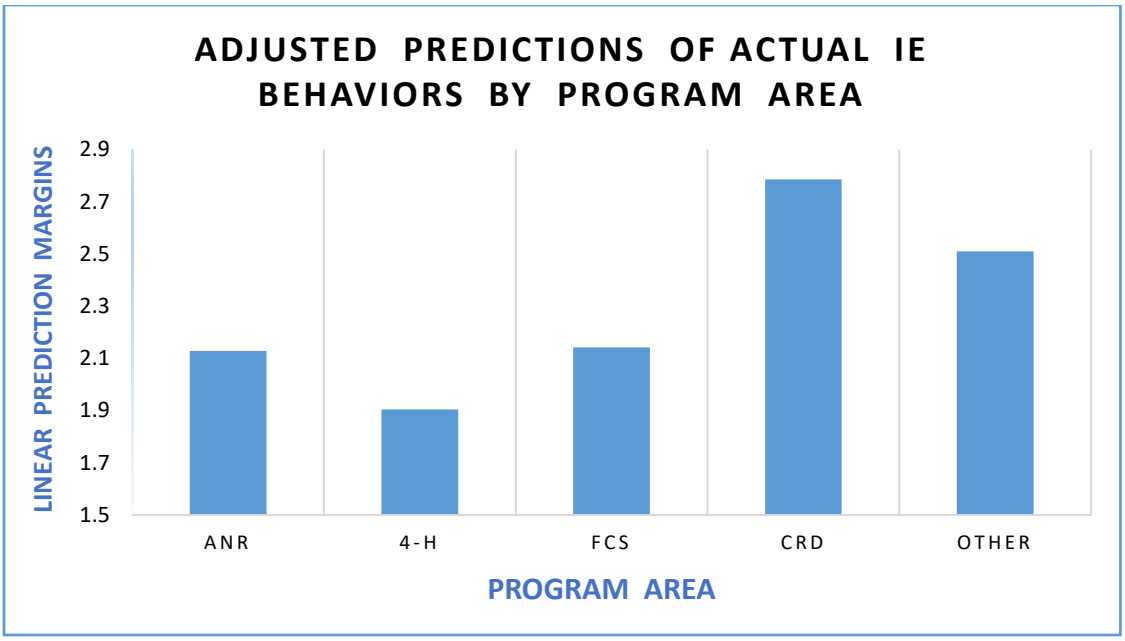

Figure 4-5. Differences in actual impact evaluation behaviors between program areas.

A one-way ANOVA conducted to determine if the effect of position type on actual IE behaviors was different for different position categories. There was a statistically significant difference between groups as reported in the ANOVA output $F(4,333) = 12.91$, $p \leq .001$. Study participants were classified into five position types a) County Agent/Extension Educator/ Program Coordinator (group 1, $n = 204$), b) Program Assistant (4-H,etc., $n = 4$), c) Specialist (state or regional, $n = 91$), d) Administrator (district, state, dept., $n = 32$), and e) Other ($n= 4$).

A Tukey post hoc test indicated that actual impact evaluation (IE) behaviors were statistically significantly higher in the specialist category compared to the county agent/Ext educator/program coordinator category (.39 ± .08 actual IE behaviors, $p \leq .001$). Actual IE behaviors were statistically significantly higher in the administrator category compared to the county agent/Ext. educator/program coordinator (.76 ± .12 actual IE behaviors, $p \leq .001$). Actual IE behaviors were also found to be significantly higher in the administrator category than in the specialist category (.43 ± .13 actual IE behaviors, $p \leq .05$). Actual IE behaviors were also found

to be significantly lower in the other category compared to the administrator category (-1.18 ± .34 actual IE behaviors, $p \leq .05$). However, there were no statistically significant differences between the program assistant and county agent/educator/program coordinator categories (-.16 ± .32 actual IE behaviors, $p = .986$, between the other and county agent/educator/program coordinator categories (-.41 ± .32 actual IE behaviors, $p = .703$), between the specialist and program assistant categories (.50 ± .33 actual IE behaviors, $p = .540$), between the administrator and program assistant categories (.93 ± .34 actual IE behaviors, $p = .052$), between the other and program assistant categories (-.25 ± .45 actual IE behaviors, $p = .982$), or between the other and specialist categories (-.75 ± .33 actual IE behaviors, $p = .148$).

Figure 4-6 Graphically represents the predictions of actual IE behaviors by position type.

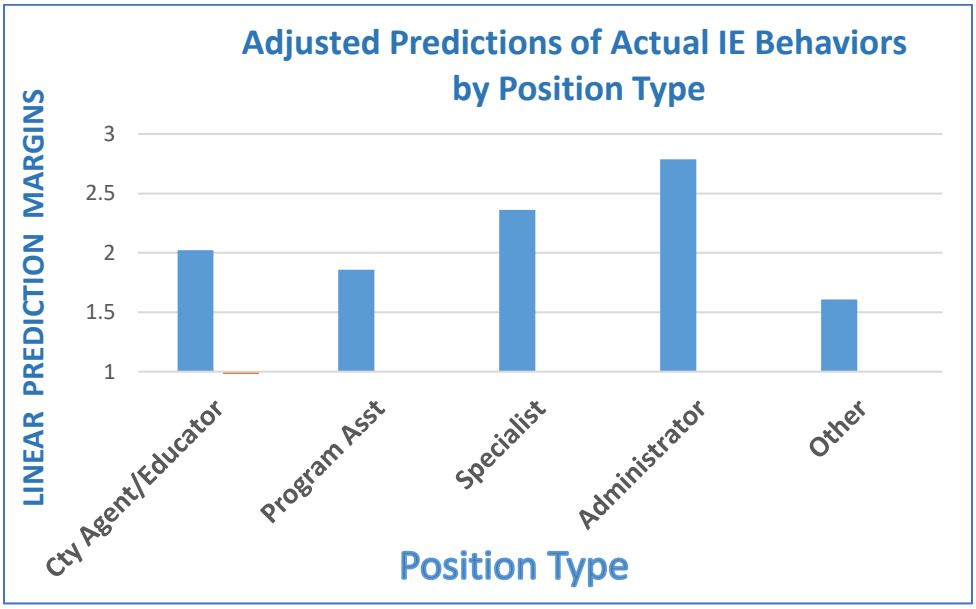

Figure 4-6. Differences in actual impact evaluation behaviors between position types.

The ANOVA comparing location/state and actual impact evaluation behaviors was not significant, indicating that there was no statistically significant relationship between the two

variables $F(4,333) = 12.91$, $p \leq .001$. The same was true for the ANOVA comparing separate evaluation department and actual impact evaluation behaviors $F(4,333) = 12.91$, $p \leq .001$.

ANOVA analysis of data from categorical variables not included in the SEM analysis showed that program area and position type were statistically significant with respect to group effects on the dependent variable of actual IE behaviors. This significance may explain some of the disturbances in the SEM model given that the factors remained external to the final model.

Qualitative Data Analysis

Two items included in the survey instrument were designed to provide qualitative data to be analyzed to identify themes that related to the focus of each question. The two open-ended were:

Q3.11. What do you need, if anything, to help you do a more effective job of completing program impact evaluations? Please describe/elaborate. _____

Q3.12. What do you see as the biggest challenges or obstacles to improving impact evaluation of programs?

Thematic analysis of the responses for each question resulted in identification of factors that may illuminate the relationships found among the variables through the SEM analysis. Themes found for each question are listed in the order of most often mentioned in Table 4-17.

The top themes are not surprising. Impact evaluation involves measuring and documenting change which has not traditionally been a focus of evaluation in Extension. The mandate to report program impact at the Federal level, based on legislation, was adopted over two decades ago. However, focused education and training in how to do effective impact evaluation has been inconsistent at the Federal and state levels. The need to do impact evaluation

has not always been communicated to county and regional educators or to all state and regional content specialists.

Table 4-17

Themes Identified in Qualitative Data from Two Open-ended Items

Themes by Overall Rank	**Related Study Factor(s)**	**Ranks of Themes**	
		Q3. 11	& Q3.12
n = 284 (85% of study *N*)		**Needs**	**Challenges/ Obstacles**
1 – How to Measure Change	Competency – Skills & Importance Training, Support	1 (23)*	1 (27)
2 – Training (general)	Competency – skill	2 (21)	3 (21)
3 – Planning for Impact	Competency – skills & importance Behavioral Intent	4 (19)	2 (22)
4 – Conflicting Priorities	Norm, Culture, Support, Number of Roles	3 (20)	4 (18)
5 – Time / timing	Culture, Norm, Attitude, Support	6 (16)	5 (16)
6 – Moving Target	Norm, Culture, Competency-Management	5 (17)	6 (14)
7 – Diverseness in: - Prog Planning - Context – priorities - Competencies - Schedules	Implicit in Extension Contexts, Norms, ProgPlanning, IE Competencies, and IE Management	7 (14)	7 (12)
8 – Resources – lacking	Support, Culture, Norm, External	8 (11)	8 (11)
9 – Espoused vs Enacted Values	Norms, Attitude, Culture, Support	Underlying Observed Theme (8, 9)	

Note: Frequencies are provided in parentheses(x)* to the right of the ranked themes for questions 3.11 and 3.12; the frequencies indicate the number of times the factor was mentioned in the responses to the question.

Understanding how to measure and plan for impact (change) were top themes identified in responses to both questions. The lack of knowledge and skills in impact evaluation planning, design, and implementation were also found to be significant themes in the quantitative data.

Timing and lack of time were also identified as top issues mentioned by respondents. These themes pervade several of the factors included in this study including culture, norms, attitude, behavioral intentions and actual performance of impact evaluation. Lack of time to do impact evaluation and lack of time for planning were mentioned most often. Time was also an underlying factor in conflicting priorities, with respondents noting that sufficient time to perform all tasks including IE and others was lacking. This also related to scarcity of resources.

The themes of conflicting priorities (#4) and moving target (#6) are related; however, conflicting priorities was most often mentioned as an issue for both individuals and for the organization. This finding relates to leadership, norms and priority management. The moving target theme was tied to the changes in requirements at both the state and the Federal level. This theme relates to organizational priorities and management and to both internal and external factors that affect both individuals and the organization.

One theme not mentioned elsewhere was that of diverseness. Respondents mentioned diversity of programming, impact evaluation competencies/skills, schedules and priorities. This may be, in part, due to the nature of Extension, which puts local *and* state educators in a position that builds conflicting priorities into their positions. Professionals at both the state and local levels are responsible to more than one entity. At the local level staff report to both local stakeholders and to their Extension supervisors. Specialists at regional and state levels are responsible to both their academic department and to Extension administrators.

Another key finding of the analysis of this data was that educational programs are not planned to produce the types of impact that is required by state and Federal reporting. Respondents mentioned that they need time, knowledge, and support to be able to plan programs that target measurable impact.

The existence of organizational norms in the established culture of Extension includes evaluation as a priority. That priority may not extend to impact evaluation in actual practice. Espoused values versus enacted, actual practice of those values is a theme that resonated with many individuals. Thus, although evaluation is touted as an essential component of Extension practice, it may not be valued highly by some. This, in turn, can affect Extension professionals' intent to do impact evaluation and their actual performance of related IE behaviors.

Results will be reviewed and discussed further to focus study findings in chapter five.

Brief Review of Qualitative Data Analysis Approach

Responses to the two open-ended questions came from the same number of respondents although of examination of the data showed that some respondents did not answer both questions. Responses were obtained from 85% of the individuals in the study sample (284 of 336); this represented a significant portion of data in comparison with the quantitative data.

Qualitative data obtained from each question were analyzed separately using thematic coding, versus coding, and multiple coding to identify themes. As coding progressed, it became apparent that most of the responses could be associated with the major constructs and factors chosen to include in the study. Thus, data were regrouped, and further analysis was done to determine the frequency of the themes and subthemes found in the data. This quantifying of the qualitative data was done to show reasons behind the quantitative results, either in support of, or in contrast to those results. Data from both questions was combined to identify common themes.

Additionally, themes not included as constructs or factors in the a priori conceptual model for the study were identified in the responses to this question. This information is helpful and may help to explain error in the SEM path analysis results. Relevant themes and pertinent quotes are included in the discussion of these results in chapter five.

Summary

The purpose and theoretical framework for the study were reviewed in the first section of this chapter. Constructs and factors measured using an online survey were described to show the connection between the theoretical framework and the factors to be examined through the methodology of structural equation modeling (SEM) which guided the study.

Descriptive statistics regarding study participant data were presented, as were SEM and qualitative data analysis results. Results were shared using appropriate tables and figures.

Chapter four provided an explanation of the structural equation modeling (SEM) statistical analyses performed; presented the structural model estimated and found to be of best fit for the theoretical/conceptual model and data collected; and reported the SEM findings.

Significant findings of the SEM analysis included the identification of the most influential factors on the dependent variable of actual impact evaluation (IE) behaviors. These were competency by perceived skill level, behavioral intention, number of roles in IE, education level, and attitude. Factors exhibiting noteworthy influence on factors other than the dependent variable within the model included training, culture, and proportion of teamwork.

Thematic analyses of the qualitative data were also reported as well as descriptive statistics and supplemental ANOVA (multiple regression) analyses. Qualitative data themes found to be most frequently mentioned by participants included the need to know how to measure change; a need for training in general; a need to plan programs to achieve impact; existence of conflicting priorities, lack of time and timing/coordination concerns; changing expectations (“moving target”) regarding impact evaluation practices and goals; and diversity in program planning, contexts, schedules, and programming.

Two categorical, exogenous variables not able to be included in the SEM model analysis were analyzed using ANOVA and were found to show statistical significance among the categories and their influences on the dependent variable of actual IE behaviors—program area and position type. Two other factors—separate evaluation unit and location (state) were not found to have statistically significant effects on the dependent variable of actual IE behaviors. These results served to provide more depth and richness of data to help interpret the overall findings of the study.

Chapter five provides a discussion of the study findings and comparisons to previous research, as well as an outline of contributions of this study to theory. Implications for application and practice are shared and recommendations for future research made based on these findings are also presented in chapter five.

CHAPTER FIVE. DISCUSSION AND CONCLUSIONS

This research was a quantitative, non-experimental correlational study of the attitudes, evaluation competencies, evaluation culture, evaluation behaviors, and individual and organizational demographic factors associated with Extension educators' impact evaluation practices in the North Central Region of the Cooperative Extension service.

Results are interpreted and implications discussed in this chapter through the lens of the concept model created for this research. Relationships of significance among the variables are examined and discussed. Qualitative results are included when appropriate to provide additional insight. Contributions of findings to theory include those related to impact evaluation, role theory, capacity building theory, and others. Implications for application and practice in the context of Extension and, possibly for similar non-profit government organizations, are discussed. Suggestions for future research are also provided.

Research Purpose and Research Question

The purpose of the study was to employ a comprehensive, systems approach to study the interrelated individual and organizational factors that affect the practice of impact evaluation (IE) in Extension. The research question explored in this study was: What are the relationships among the individual and organizational contextual factors that influence impact evaluation behaviors within Cooperative Extension? Constructs and factors explored in the study are shown in Figure 4-1.

Theoretical and Conceptual Foundations of the Study

A major feature of this research design was the decision to use the Structural Equation Modeling (SEM) statistical methodology and analysis technique. The use of SEM to organize and implement this study facilitated simultaneous exploration of many constructs theorized to be involved in impact evaluation behaviors in the context of Extension educational program evaluation. The theoretical/conceptual model created for this study was presented in Figure 1-1.

The adapted initial structural model shown in Figure 5-1 includes the original variables of Ajzen's Theory of Planned Behavior (TPB) behavioral analysis model plus three additional constructs: a) competencies-skill, b) competencies-importance, and c) IE culture. The model also includes both individual and organizational demographic variables.

Theories incorporated into the conceptual model for the study, in addition to the TPB, included organizational culture theory, learning organization theory, role theory, as well as competency and capacity building literature. Other contributing literature included research and professional journal articles regarding evaluation in general and impact evaluation in Extension (Lamm, 2011; Ghimire & Martin, 2011; McClure, et al., 2012; Wells-Marshall, 2012; Workman, 2010; and Vengrin, 2016), the experience of the researcher in three states in the North Central Region, and guidance from Extension experts and administrators involved in impact evaluation (IE).

The initial structural model specified after data collection according to SEM guidelines is depicted in Figure 5-1. The model includes the original variables of attitude, subjective norm, perceived behavioral control, behavioral intent, and actual behaviors which are the core of Ajzen's Theory of Planned Behavior (TPB) behavioral analysis model; these foundational factors are shown in blue in the model. Three additional endogenous factors/constructs completed the core of the expanded TBP model: a) competencies by perceived skill level, b) competencies by perceived importance, and c) IE culture and are shown in green. The model also includes both individual and organizational demographic variables shown in salmon in the far-left portion of the model.

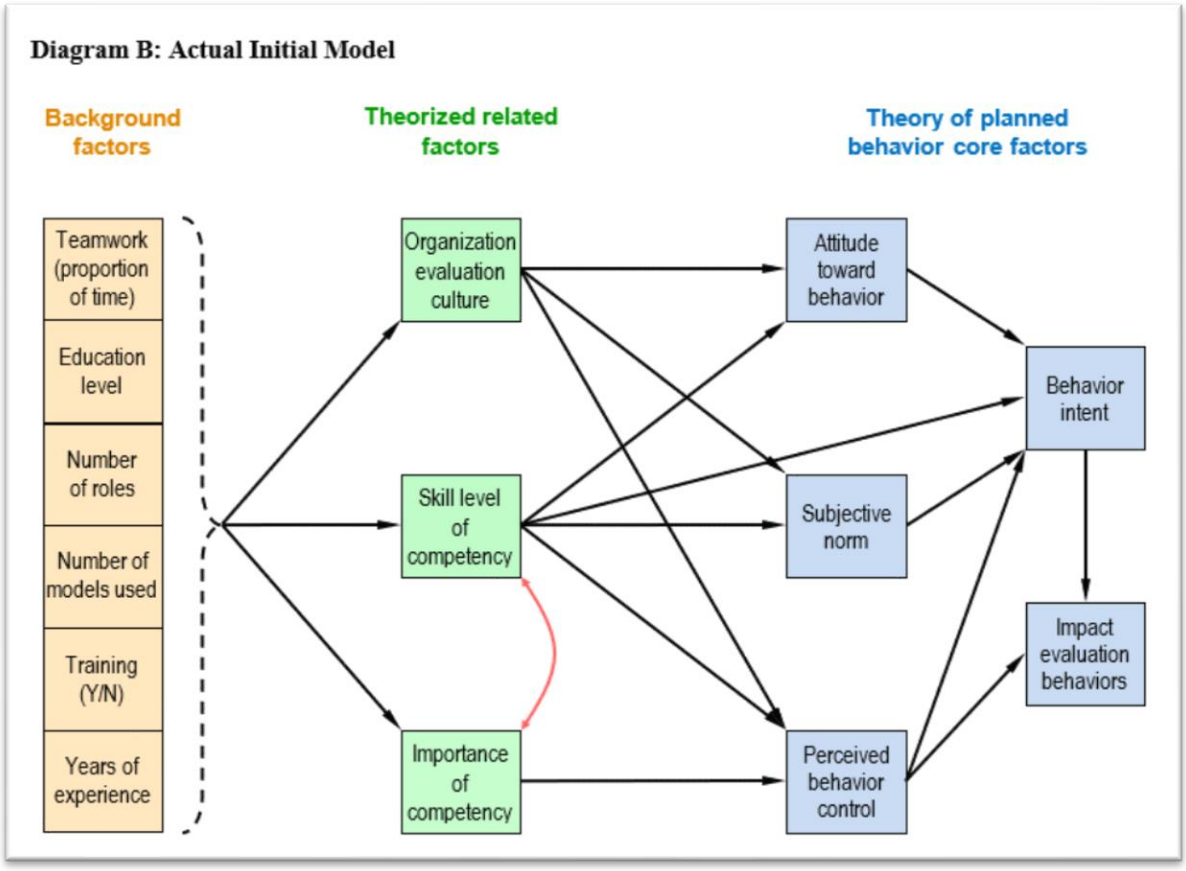

Figure 5-1. Initial model. Black arrows represent direct effects (pathways). The red double-headed arrow indicates correlated disturbance (error) terms.

Impact evaluation planning, capacity building, and implementation are complex, interdependent, multi-faceted and dynamic and may best be understood using a comprehensive, systems approach (Urban, Burgemaster, Archibald, & Byrne, 2015; Trochim, Urban, Hargraves, Hebbard, Buckley, Archibald, & Burgemaster, 2012). Behavioral research has shown that an individual's attitudes, beliefs, and expectations are important predictors of behavior (Ajzen & Fishbein, 1980; Ajzen, 1991; Bandura, 1977).

One must also consider the effects of the context in which an individual is embedded regarding a specific target behavior such as impact evaluation. Context includes organizational culture and subjective norms; internal and external influences on individuals; and perceptions of organizational support for impact evaluation. Stame and Presti (2017) posited that evaluation

strategies which focus on positive perspectives and success contribute to organizational learning. Collaboration and teamwork have been cited in evaluation research studies as essential to the effective design, implementation, and evaluation of educational intervention programs (Lamm, 2011; O'Sullivan, 2012; Patton, 2008; Preskill & Boyle, 2008; Roche, 1999). Research on the roles of evaluators in Extension is limited and has not included consideration of the number of roles involved in actual evaluation behaviors. Research on the number of IE roles performed was not found in the literature; the factor was included based on role theory research completed in other contexts.

Review and Discussion of the Major Findings of the Study

Results of the SEM analysis are represented in the final structural model found to best fit the data which is depicted in Figure 5-2 below.

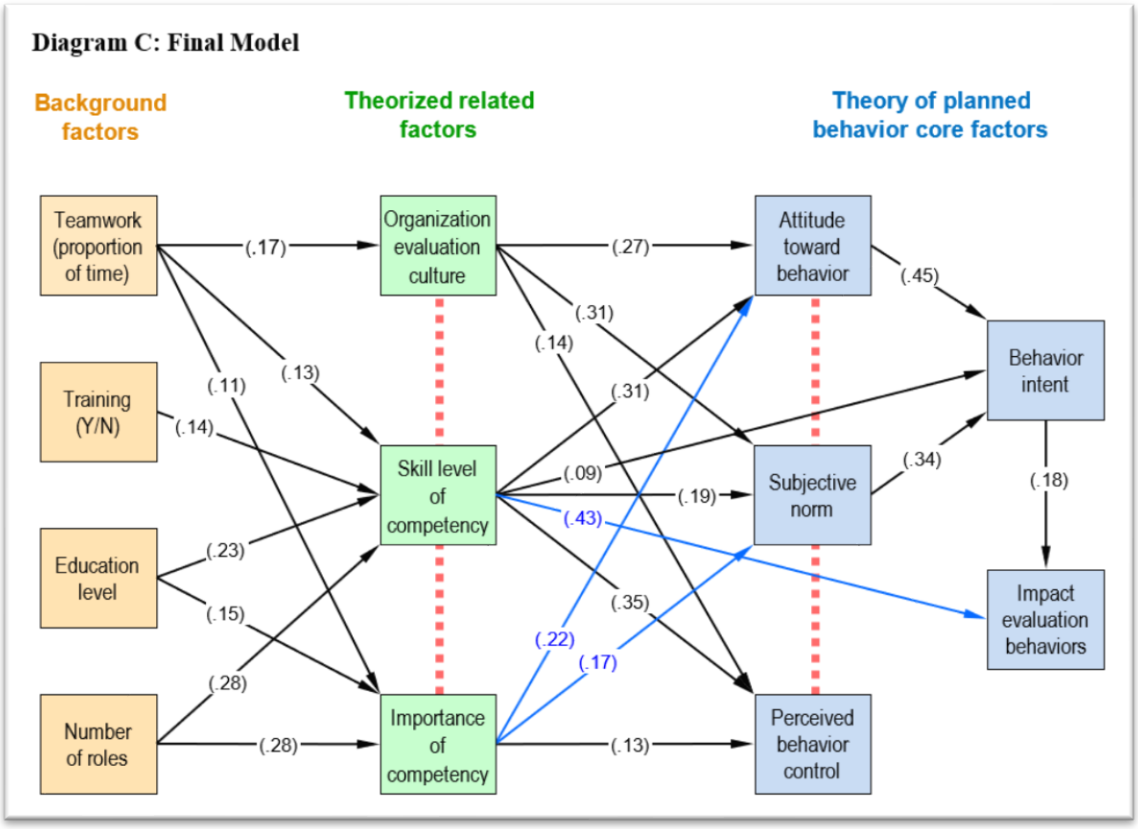

Figure 5-2. Final model after modification. Black arrows represent direct effects (pathways). The blue paths were added to the model to achieve acceptable fit. The standardized path coefficient estimates are shown in parentheses. The red dashed lines are used to indicate subsets of endogenous variables that have correlated disturbance (error) terms.

The most influential factors in the SEM structural model were found to be competency by skill level with the highest total effect (β = .49) and the most direct effects to other factors in the model, making it the primary influence over the dependent variable. Behavioral intent was the second most influential factor in the model with one direct effect direct (β = .18) on actual impact evaluation behaviors. The third highest total effect in the model was shown by the exogenous factor of number of roles, which had a small effect (β = .14) on the DV of actual IE behaviors. Number of roles had small indirect effects on three other endogenous variables within the model: intent (β = .13), attitude (β = .15), and subjective norm (β = .07).

The fourth highest total effect and level of influence in the model was the factor of education level, followed by attitude toward the DV of actual impact evaluation behaviors, and training. Proportion of teamwork had a small total effect (β = .028), within the model although it exhibited direct effects on three other factors and indirect effects on five other factors. Teamwork was significant in the model although it did not have a strong total effect. Teamwork and collaboration were cited repeatedly in evaluation and impact evaluation literature as essential to the effective design, implementation, and evaluation of educational intervention programs (Lamm, 2011; O'Sullivan, 2012; Patton, 2008; Preskill & Boyle, 2008; Roche, 1999).

The finding that competency by skills was most influential in the model is supported by literature on competency and capacity building (Preskill and Torres, 1999).

This study showed that the factors of location and position type exhibited different variation in effects on the dependent variable of actual impact evaluation behaviors. Analysis was done using ANOVA because the data were categorical and were not included in the SEM path analysis. Berrio (1999) found differences in organizational learning and culture within the different program areas in Extension. Ghimire and Martin (2013) also found that needs for training and staff development in evaluation in general varied by program area.

The factor of position type was not included in the SEM model estimation. It was found to be significant through ANOVA analysis, however. This finding is inconsistent with that of Cousins et al. (2014) who also observed no effects of position on capacity to do and use evaluation for internal evaluators in an organization.

A small direct effect (β = .17) was found to exist between the factor of proportion of teamwork and IE evaluation culture. Proportion of teamwork also had small direct effects on both competency by perceived skill (β =.13) and competency by perceived importance (β =.11). The overall total effect on the dependent variable of actual IE behavior was significant All effects noted here were significant within the model. The findings that teamwork clearly influences impact evaluation behaviors is significant and contributes to knowledge about the strategic use of teamwork to support IE behaviors in Extension.

In the current study the effects of culture on the dependent variable of actual IE behaviors were mediated by norm, by attitude, and by intention. This is congruent with findings by Smith (2015), Wells-Marshall (2012), and Vengrin who found that those factors related to attitude, norms, and culture affected performance of education evaluation.

The factor of competency by skill showed strong effects on both intent and actual behavior, thus becoming more significant in the final model estimated for this study. Competency by skill also had a medium direct effect (β = .35) on perceived behavior control which was not surprising. The possibility that perceived behavior control (pbc) was affected by the fact that impact evaluation is mandated in Extension may have stifled connections between pbc and both intent and actual IE behaviors. It is also prudent to note that there were disturbance correlations between pbc and both subjective norm and attitude, which shows that the factors covary and may partially explain the lack of direct influence on intent and actual IE behaviors.

Findings of ANOVA analysis of data regarding program area for this research supports the findings of studies conducted by McClure, Fuhrman, and Morgan (2012) in Georgia, by Ghimire and Martin in eleven states in the North Central region of Extension in 2013, and by Ghimire and Trechter (2012) in Wisconsin. These studies found that evaluation competence of Extension educators differ by their area of program responsibility.

The mediating variables described here that have indirect effects show various levels of influence on the desired impact evaluation behaviors. These mediating variables help to explain why a target behavior may be performed or not.

Ajzen (1989) noted that perceived behavior control may not be a particularly useful indicator of behavioral intention or actual target behaviors if individuals have "relatively little information about the behavior, when requirements of available resources have change, or when new and unfamiliar elements have entered the situation," (p. 251). This may explain why the factor of perceived behavior control, although significant in the model, was not found to have direct causal relationships to either intention or actual impact evaluation behaviors in this study. Perceived behavioral control was directly affected by the factors of culture, competency by perceived skill level, and competency by perceived level of importance.

Findings of this study showed that, while study participants rated themselves competent in some of the various skills needed to perform impact evaluation, responses to the open-ended questions showed that many respondents considered their biggest challenge was to know "how to measure change." Further, participants from one of the six states represented in the study noted that there was a major restructuring of their evaluation support unit in progress. Thus, both the information about the behavior and uncertainty about available support resources may have affected the perceived behavioral control of participants. Ajzen further noted that a significant

path from “perceived behavioral control to behavior is expected to emerge only when there is some agreement between perceptions of control and the person’s actual control over the behavior,” (p. 251).

Given that the structural model explains only a portion of the variance in the target behavior of impact evaluation behaviors (28.6%), it is clear that there are other factors of influence not included in the model. The “missing” factors most likely include the categorical variables representing factors included in the original a priori specified model that could not be included in the SEM analysis of the model but were found to be significant through ANOVA analysis. Those factors were type of position and program area. Other factors not included were years of experience, support and program planning models.

The factor of support was also omitted from the model due to lack of reliability of the measure. The researcher recommends reintroducing that factor back into the model and measuring it using a Likert scale to assess the perceived level of support available in the organizational environment. A second way to assess support would be through determining if a participating organization has a separate evaluation unit, which was the goal of the item, through conversations with administrators. An additional way to assess support would be to determine if there are evaluation experts or specialists available within the organization. The level of support could be quantified as: a) no separate unit or specialists = 0; b) a separate unit or specialists = 1; c) both a support unit and specialists = 3. This would produce a continuous variable which could be included in SEM analysis of a specified model.

Discussion of the Results of the Qualitative Data Analysis

Understanding how to measure and plan for impact were top themes identified in the qualitative data. The lack of knowledge and skills in impact evaluation planning, design, and

implementation were also found to be issues in the quantitative data. Thus, the qualitative and quantitative data support the findings that knowledge of methodology and planning for impact evaluation are challenges for Extension professionals throughout the region addressed by participants in the study. Regardless of the fact that individuals have had training in impact evaluation, many still feel their competency for performing effective impact evaluation is limited. Many study participants gave responses similar to the following comment: "I don't have knowledge of research methods needed for IE [impact evaluation]." Waddell (1991) and others (Wippersberg, 2017; Roche, 1999) have expressed the concern that evaluation has become more akin to research and evaluation research in the past two decades due to a focus on quantitative evidence by government and other stakeholders who require impact evaluation. If Extension professionals are to do an effective job of impact evaluation, do they need training and staff development in more rigorous research methods? Do internal evaluators who are primarily educational program presenters need to have that depth and breadth of knowledge? These questions pose challenges for Extension and other non-profit and government organizations who are required to compete evidence-based impact evaluation aimed at demonstrating that change has occurred due to an educational program intervention. There may be a need to hire more evaluation experts within the Extension organization to serve as team leaders and to oversee impact evaluation processes.

Timing and lack of time were also identified as top issues mentioned by respondents in the qualitative data. These themes pervade several the factors included in this study including culture, norms, attitude, behavioral intentions and actual performance of impact evaluation. Lack of time to do impact evaluation and lack of time for planning were mentioned most often. Time was also an underlying factor in conflicting priorities.

The themes of conflicting priorities and "moving target" are related; however, conflicting priorities was most often mentioned as an issue for both individuals and for the organization. These findings relate to leadership, norms and priority management. The moving target theme was tied to changes in requirements at both the state and the Federal level by respondents. These concerns relate to internal and external factors that affect both individuals and the organization.

Another key finding of the analysis of the qualitative data was that educational programs are not planned to produce the types of impact that is required by state and Federal reporting. Respondents mentioned that they need time, knowledge, and support to be able to plan programs that target measurable impact.

Anther related theme found in the qualitative data that was not mentioned elsewhere was that of diversity. Respondents mentioned diversity of programming, skills, schedules and priorities. This may be, in part, due to the nature of Extension, which puts local *and* state educators in a position that builds conflicting priorities into their positions. Professionals at both the state and local levels are responsible to more than one entity. At the local level staff report to both local stakeholders and to their Extension supervisors. Specialists at regional and state levels are responsible to both their academic department and to Extension administrators. Further, if educational programming is planned with local and regional needs in mind, rather than or in addition to Federally driven "top-down" programming, the diversity in planning for impact producing programming which meets local needs will increase. This has implications for practice, and capacity building, as well.

Some respondents noted that they did not have the level of expertise in research needed to perform adequate impact evaluation. This is a valid concern for those in Extension. Picciotto (2014) noted that current approaches to impact evaluation tend to be more like research than

traditional evaluation. Focusing on the effects that can be validly attributed to an intervention education program is key to successful impact evaluation. Many programs are treated as experiments and reduce impact evaluation to one concern: did the intervention work as intended? (p. 32). This study showed that Extension professionals need more assistance in providing evidence to answer that question as well as to plan better for intended impacts.

The existence of organizational norms in the established culture of Extension includes evaluation as a priority. That priority may not extend to impact evaluation in actual practice. Espoused values versus enacted, actual practice of those values is a theme that resonated with many individuals. One individual expressed this concern in this way: "Although we are told that we must do impact evaluation of programs, leadership doesn't adequately encourage, support or guide educators and teams to conduct impact worthy programming."

Theoretical Contributions

This section will discuss and relate study findings to previous research and make connections to related theory.

The finding that competency by perceived skill level was a strong factor in predicting both behavioral intention and the actual performance of impact evaluation (IE) behaviors is significant. The competencies included in this study included skills, attitudes, and knowledge in the domains specified by the American Evaluation Association (AEA) in the areas of a) professional practice, b) methods, c) context, d) planning and management and e) interpersonal skills. These competency domains are similar to those established by Canada, Australia, and the International Board of Standards for Training, Performance and Instruction. The finding that competency by perceived skills level influences a number of variables in the final SEM model contributes to evaluation theory in the area of conditions for effective practice.

Caution needs to be exercised in the generalization of the results of this study given that Ajzen and others noted that, beyond trying to ensure that people have accurate information to perform a desired behavior, it is important to determine what information they already have and how it interacts with intentions and actual behaviors, regardless of how accurate and important the information seems (Ajzen, Joyce, Sheikh, & Cote, 2011). Concern must not be for general information but for specific information and knowledge that guide intent and result in actual IE behavior performance. Thus, results of the study also have implications for competency/capacity building theory and strategies which are guided by that theory.

The finding that, as number of roles increased, performance of actual IE behaviors also increased is congruent with role theory research regarding the effects of participating in multiple roles which shows that positive results are more likely than negative (Biddle, 2013; Marks & MacDermid, 1996; Van der Horst, 2016). Marks (1977) posited that level of commitment affects and is affected by multiple role participation. However, the findings from the qualitative data somewhat contradict this finding, as lack of time and conflicting priorities were cited as challenges in performing impact evaluation. This may indicate that, when other roles, beyond those involved with impact evaluation, take priority, the result may be temporary role overload, which is defined as "a time-based form of role conflict in which an individual perceives that the collective demands imposed by multiple roles exceeds available time and energy resources, thereby making an individual unable to adequately fulfill the requirements of the various roles," (Higgins, Duxbury, & Lyons, 2010; p. 847-848).

The finding that teamwork clearly influences impact evaluation behaviors is significant and contributes to knowledge about the strategic use of teamwork to support IE behaviors in Extension. Team theory and influences on impact evaluation behaviors

The primary finding in the qualitative data was that study participants felt the need to learn "how to measure change" and a secondary need for more training in IE in general. Lamm, Israel, and Diehl (2013) had similar findings and noted that "some Extension professionals may lack the expertise to perform evaluation behaviors that measure long term change or conduct advanced inferential statistics," (p. 6).

The findings in the qualitative data that there is a need for customized impact evaluation designs and that a lack of time to deal with IE and evaluation in general exists relate directly to adult learning theory which stipulates that adults wish to learn what is relevant to them in their current situation and at a convenient time to be put into practice when needed (Knowles, 1990; Knowles, Holton, & Swanson, 1998).

Finally, the factor of diverseness of context, schedules, and program planning underscores the importance of context to doing effective IE. This finding concurs with the acknowledgment of diverse experiences as significant to the success of educational intervention programs in the Handbook on impact Evaluation by The World Bank (Kandker, Koolwal & Samad, 2009).

The need to plan for impact before attempting to measure it underscores the importance of good planning as well as an understanding of basic program planning models. These findings align with program theory evaluation planning (Rogers, Petrosino, Hueber, & Hacsi, 2000). Making program goals and evaluation plans explicit can help increase the possibility of documenting evidence of program impacts. Related evaluation capacity building is consistent with the tenets of basic learning theory that measurable objectives are needed to guide and ensure that educational efforts are successful.

Implications for Future Research

For future research, this study raised the issue of diversity of competencies, attitudes, and expectations regarding IE. It is recommended that future research take a closer look at the types of human support and training needed by those in various roles such as data analyzers, IE designers, and others.

Asking what works well would also be a prudent research approach to determine whether findings of this and other research regarding evaluation and IE in Extension are relevant to actual practice in Extension impact evaluation. Rather than approaching needs assessment in a scarcity mode, it may be best to ask what is needed or presents a challenge and then focus on what actually works to determine more practical, doable strategies for Extension impact evaluation. Silliman, Crinion, and Archibald (2016a) studied evaluation champions within Extension and found that they exist in most states and may serve as a valuable resource if encouraged and nurtured by program evaluation experts.

Mining the data collected in this study to do more explorative research and/or repeating this type of study in other non-profits and/or Extension is recommended to examine the different roles that professionals fill in the impact evaluation process. This study introduced the concept of roles to be filled as the first question in the survey. This served to heighten participant's awareness of the different roles involved in effective impact evaluation of educational programs. Exploring the number of roles filled was the focus of this research, based on role theory which stipulates that filling multiple roles has a more positive effect on individuals involved in a process. Deconstructing the list of roles to determine who fills the roles, if their specific role(s) correlate with attitudes regarding impact evaluation, and identifying other factors which may have a relationship to specific roles are just a few of the approaches that might be taken.

Research into the perceived levels of organizational support for impact evaluation is also suggested. This study attempted to determine if participants whose location/state had a separate department or unit that focused on program planning and evaluation was an influential factor in completing impact evaluations in their contexts. Participants answered the question of whether their state had a separate evaluation unit yes or no and, when data were examined, it was found that some from states that did not have such units answered "yes" while some in states with such a unit answered "no;" thus, the question was judged to have low internal reliability and removed from the model. The researcher wondered if, as in some cases, perception might be reality and, thus, yield valid results; however, that idea was put aside in favor of reliability. Future research could ask if participants feel that they have adequate support in their state and answer on a Likert scale of strongly disagree to strongly agree to provide continuous data for analysis in a SEM model. This approach could yield results that might assist administrators and others in their planning to support and catalyze impact evaluation in their organizations.

Use of SEM Methodology to Frame and Conduct Research

The complexity of the interactions among the parameters identified for exploration in this study are difficult to unravel through linear regressions alone, as cited in previous research (Lamm, 2011; Vengrin, 2016). Structural equation modeling may be used to determine which individual and organizational factors directly and indirectly influence target behaviors. Studies accomplished using other research methods may be repeated using SEM to confirm factors in an effective model to explain whether successful impact evaluation is accomplished. Such models could be of use to those planning for support for professionals through designing interventions to increase positive attitudes, intentions, and values regarding impact evaluation. Thus, use of the SEM methodology is recommended for future research regarding impact evaluation systems.

Implications for Practice

This study and others that focus on understanding evaluation behaviors and practices, organizational culture, and related factors within Extension serve to increase awareness of the fact the organization has a culture of evaluation, whether it is recognized or not. Further, this study focused on factors that influence impact evaluation behaviors and practices, specifically. There may be differences between an organization's overall evaluation culture and the more specific subculture related to impact evaluation. Examining the relationships among the factors that affect impact evaluation , specifically, may help Extension professionals enhance impact evaluation through targeted, intentional capacity building efforts aimed at those relationships identified in the study.

The finding that many Extension professionals perceive a need to learn how to measure change has implications for professional development for Extension administrators, specialists, county- and regionally-based professionals. Professionals at all levels indicated a need to learn how to measure change in their responses to the open-ended questions regarding needs and challenges related to performing impact evaluation behaviors. Intentional capacity building efforts that focus on skill sets specific to the measure of medium- and long-term change are recommended.

Assessment of competencies needed, using the Borich method to determine mean weighted discrepancy scores to indicate gaps in competencies based on employees' self-rated skills and perceived importance of current competencies is recommended. The use of the Borich method of needs assessment is recommended to customize the design of capacity building efforts for Extension professionals in specific states. Furthermore, focusing on particular programs may also help create the types of impact evaluation designs that can be useful in the context of

locations that are implementing common programs. This approach may also provide opportunities for comparison of impact evaluation data, encourage collaboration and teamwork among those assessing the impact of similar programs, and help build impact evaluation capacity for both individuals and the organization.

The correlation between the factors of attitude and subjective norm found in this study are congruent with the findings of other researchers (Armitage & Conner, 2001; Vengrin, 2016). Those researchers noted that both attitude and subjective norm are closely associated with each other and with evaluation culture; thus, they influence overall behavioral processes (Armitage & Conner, 2001; Conner & Armitage, 1998; Vengrin, 2016). Implications for practice include working to intentionally affect attitudes, norms, and related evaluation culture to ensure a more positive context for impact evaluation behaviors.

Encouraging teamwork and collaboration is another recommendation based on the results of this study which found that the proportion of impact evaluation done through teamwork correlated positively with Extension professionals' actual performance of impact evaluation.

The findings that individuals feel they need to know how to measure change, that they receive mixed messages due to conflicting priorities, that programs are not planned with impact as a goal, and that impact evaluation goals are seen as "moving targets" has implications for Extension impact evaluation leaders. The creation of consistent, intentional, holistic approaches to evaluation capacity building to support impact evaluation behaviors is recommended. Working to make the planning of IE an integral part of program planning through use of program logic process-based planning and focusing on communicating consistent IE goals and processes to Extension professionals can help create a more positive environment for all involved in IE.

Limitations

The use of SEM analysis to provide answers to the research question may limit generalizability, according to some (Brannick, 2019; Kline, 2016). The SEM methodology is also considered to be in the relative early development stages for a statistical technique which may limit the understanding of its potential within the ranks of academicians as well as possibly being a barrier to the publishing of a study's results.

A potential threat to the external validity of this study is the fact that all Extension professionals who participated were from one geographical region of the United States which may affect the generalizability of the study results. Furthermore, data collection was done via a self-reported online survey which has the potential for response bias. Care was taken to ensure that techniques such as using Likert-scaled items to give participants the option to answer as accurately as possible by using a scale of "Strongly disagree", "Disagree", "Neutral", "Agree" and "Strongly agree." Coding response options into an interval scale of 1-5 can help you a researcher obtain more accurate and effective answers (DeVellis, 2016).

The ANOVA analysis performed to determine if different program area categories affected the dependent variable of actual impact evaluation behaviors differently. There may be a problem with the data given that a substantial proportion of respondents selected "other" as their primary program area and the responses given were very diverse; thus, the category of "other" responses was very diverse, which may affect the validity of the ANOVA test for that category.

One final limitation may be that the data collection instrument, although it contained many items that were previously validated by other researchers, was not pilot tested. Expert review was used to increase construct validity and Cronbach's alpha results for the items in the instrument showed high internal reliability.

Summary

Results were interpreted and implications discussed in chapter five through the lens of the concept model created for this research. Relationships of significance among the variables are examined and discussed. Qualitative results are included when appropriate to provide additional insight. Results are reported and reviewed in relationship to previous research, including recommendations for future research. The researcher includes suggestions for impact evaluation practice and for proposed evaluation capacity building strategies designed to strengthen impact evaluation competencies and evaluation culture within the Extension service.

The relationships found among the factors will contribute the knowledge of Extension educators who completed the survey, those who lead evaluation capacity building efforts for impact evaluation competencies and methods, and to the academic literature in the areas of attitude-behavior relationships, professional evaluation competencies, and the understanding of Extension evaluation culture.

Discussion included review of the significant factors and the influences on the target behavior of impact evaluation in Extension. Detailed information on relationships found in the SEM analysis of the data will be reviewed such as the main effects found in path analysis. Supporting qualitative data will be tied to quantitative findings. Findings that showed contrasting views and perceptions among the participants will be reviewed, also.

Implications for practice and recommendations for future research include further exploration of influential factors specific to impact evaluation in order to design strategies for capacity building in Extension and, possibly, in related non-profit and government organizations.

REFERENCES

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In *Action Control* (pp. 11-39). Springer Berlin Heidelberg.

Ajzen, I. (1988). *Attitudes, personality and behavior.* Chicago, IL: The Dorsey Press.

Ajzen, I. (1989). Attitude structure and behavior. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *The third Ohio State University Vol. on attitudes and persuasion. Attitude structure and function* (pp. 241-274). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211.

Ajzen, I. (2002). Residual effects of past on later behavior: habituation and reasoned action perspectives. *Personality and Social Psychology Review*, *6*(2), 107–122.

Ajzen, I. (2003). Theory of planned behavior. *Social Psychology: Social Cognition and Social Perception*, 347–377.

Ajzen, I. (2006). Behavioral interventions based on the theory of planned behavior. Online: http://people.umass.edu/aizen/pdf/tpb.intervention.pdf

Ajzen, I. (2006). Constructing a theory of planned behavior questionnaire. *Unpublished Manuscript.* Online: http://people.umass.edu/~aizen/pdf/tpb.measurement.pdf

Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology & Health*, *26*(9), 1113–1127. http://doi.org/10.1080/08870446.2011.613995

Ajzen, I. (2015). The theory of planned behaviour is alive and well, and not ready to retire: a commentary on Sniehotta, Presseau, and Araújo-Soares. *Health Psychology Review*, *9*(2), 131-137.

Ajzen, I., & Driver, B. L. (1991). Prediction of leisure participation from behavioral, normative, and control beliefs: An application of the theory of planned behavior. *Leisure sciences*, *13*(3), 185-204.

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. *The handbook of attitudes*, *173*, 221.Washington, D.C.: American Association for Higher Education.

Ajzen, I., Joyce, N., Sheikh, S., & Cote, N.G. (2011). Knowledge and the prediction of behavior: The role of information accuracy in the Theory of Planned Behavior, *Basic and Applied Social Psychology, (33)*2, 101-117. doi: 10.1080/01973533.2011.568834

Ajzen, I., & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of experimental social psychology*, *22*(5), 453-474.

Alkin, M. C., Christie, C. A., & Vo, A. T. (2012). Evaluation theory. *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*, 386.

Alkin, M.C., & Taut, S. M. (2003). Unbundling evaluation use. *Studies in Educational Evaluation, 29*, l-12.

American Evaluation Association. (n.d.c). *Topical interest groups*. Retrieved from http://www.eval.org/p/cm/ld/fid=11

American Evaluation Association. (2003). *Scientifically based evaluation methods.* Online (5/11/2008) at: http://www.eval.org/doestatement.htm

American Evaluation Association. (2018). Evaluator competency domain list. Online (5/11/2008) at: http://www.eval.org/doestatement.htm

Andrews, M. (1983). Evaluation: An essential process. *Journal of Extension, 21*(5), 8–13.

Argyris, C., & Schön, D. A. (1978). Organizational learning: A theory of action perspective. Reading, MA: Addison-Wesley.

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology, 40*(4), 471–499.

Arnold, M. E. (2015). Connecting the dots: Improving Extension program planning with program umbrella models. *Journal of Human Sciences and Extension Volume, 3*(2).

Arnold, M. E. (2006). Developing evaluation capacity in extension 4-H field faculty: A framework for success. *American Journal of Evaluation, 27*(2), 257–269. doi:10.1177/1098214006287989

Arum, R., & Roksa, J. (2011). Limited learning on college campuses. *Society, 48*(3), 203.

Association of Public and Land-grant Universities. (2012). The land-grant tradition. Washington, DC: Association of Public and Land-grant Universities. Retrieved from http://www.aplu.org/library/the-land-grant-tradition/file

Astin, A. W., Banta, T. W., Cross, K. P., El-Khawas, E., Ewell, P. T., Hutchings, P., ... & Moran, E. (1992). Principles of good practice for assessing student learning. *AAHE Bulletin*, *45*(4).

Bates, R. (2004). A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. *Evaluation and program planning*, *27*(3), 341-347.

Babbie, E. (2001). *The practice of social research*. Belmont, CA: Wadsworth Publishing.

Bandura, A. (1977a). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84,* 191-215.

Bandura, A. (1977b). *Social learning theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 7*(2). 122-147.

Bandura, A. (1984). Recycling misconceptions of perceived self-efficacy. *Cognitive Therapy and Research.8,* 213-229.

Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist,* 44, 1175-1184.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50,* 248-287.

Barney, J. B. (1986). Organizational culture: can it be a source of sustained competitive advantage? *Academy of management review*, *11*(3), 656-665.

Bates, R. (2004). A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning*, *27*(3), 341-347.

Baughman, S., Arnold, M., Boyd, H. H., Franz, N. K., Mead, J. P., Rowe, E., & Silliman, B. (2010). Evaluating for impact: Professional development educational content delivery through learning communities. *Journal of Extension, 48*(3). Article 2TOT3.

Baughman, S., Boyd, H. H., & Franz, N. K. (2012). Non-formal educator use of evaluation results. *Evaluation and Program Planning, 35*(3), 329–336.

Baum, E. (2015). From frustrated to empowered: Exploring the process of how mid-level student affairs professionals make meaning of the responsibility for assessing student learning outcomes. (Unpublished doctoral dissertation). George Mason University, Fairfax, Virginia.

Bechar, S., & Mero-Jaffe, I. (2014). Who is afraid of evaluation? Ethics in evaluation research as a way to cope with excessive evaluation anxiety: insights from a case study. *American Journal of Evaluation*, *35*(3), 364-376.

Bennett, C. (1975). Up the hierarchy. *Journal of Extension, 13*(2), 7-12.

Bennett, C. (1976). Analyzing impacts of extension programs, ESC-575. Washington, D.C.: Extension Service-U.S. Department of Agriculture.

Bennett, C. (1996). Rationale for public funding of agricultural extension programs. *Journal of Agricultural and Food Information, 3*(4), 3-25.

Bennett, C., & Rockwell, K. (1995). Targeting outcomes of programs (TOP): An integrated approach to planning and evaluation. *United States Department of Agriculture: Washington*.

Berrio, A. A. (1999). *Organizational culture and organizational learning in public, non-profit institutions: A profile of Ohio State University Extension.* Unpublished doctoral dissertation. Ann Arbor, MI: The Ohio State University.

Biddle, B. J. (2013). *Role theory: Expectations, identities, and behaviors*. New York: Academic Press, Inc.

Boone, E. J., Dolan, R. J., & Shearon, R. W. (1971). *Programming in the Cooperative Extension service: A conceptual schema*. Raleigh, NC: North Carolina Agricultural Extension Service.

Borich, G. D. (1980). A needs assessment model for conducting follow-up studies. *Journal of Teacher Education, 31*(3), 39–42.

Botcheva, L., White, C. R., & Huffman, L. C. (2001). *Assessing learning culture scale*. Palo Alto, CA: Children's Health Council.

Botcheva, L., White, C. R., & Huffman, L. C. (2002). Learning culture and outcomes measurement practices in community agencies. *The American Journal of Evaluation, 23*(4), 421-434.

Brown, J. D., & Dutton, K. A. (1995). Truth and consequences: The costs and benefits of accurate self-knowledge. *Personality and Social Psychology Bulletin*, *21*(12), 1288-1296.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage.

Boyd, N. G., & Vozikis, G. S. (1994). The influence of self-efficacy on the development of entrepreneurial intentions and actions. *Entrepreneurship theory and practice*, *18*, 63-63.

Bohner, G., & Dickel, N. (2011). Attitudes and attitude change. *Annual Review of Psychology*, *62*, 391-417.

Brannen, J. M., & O'Connell, R. (2015). Data Analysis 1: Overview of data analysis strategies. In: The Oxford handbook of multimethod and mixed methods research inquiry. New York, NY: Oxford University Press.

Brannick, M. T. (2019). Structural equation modeling (SEM). Tampa, FL: University of South Florida. Online: http://faculty.cas.usf.edu/mbrannick/regression/SEM.html

Bryson, J. M., Patton, M. Q., & Bowman, R. A. (2011). Working with evaluation stakeholders: A rationale, step-wise approach and toolkit. *Evaluation and Program Planning*, *34*(1), 1-12. doi: 10.1016/j.evalprogplan.2010.07.001

Burke, W. W., & Litwin, G. H. (1992). A causal model of organizational performance and change. *Journal of Management*, 18(3), 532-545.

Buchanan, H., & Kuji-Shikatani, K. (2014). Evaluator competencies: The Canadian experience. *Canadian Journal of Program Evaluation*, *28*(3). Online: http://cjpe.journalhosting.ucalgary.ca/cjpe/index.php/cjpe/article/view/171

Carpenter, S., & Stimpson, M. T. (2007). Professionalism, scholarly practice, and professional development in student affairs. *NASPA Journal*, *44*(2), 265-284.

Charmaz, K. (2014). Constructing grounded theory (2nd ed.). Thousand Oaks, CA: Sage.

Choi, M., & Ruona, W. E. (2011). Individual readiness for organizational change and its implications for human resource and organization development. *Human Resource Development Review*, *10*(1), 46-73.

Choi, M., & Ruona, W. E. (2013). Individual readiness for organizational change. *The Wiley-Blackwell Handbook of the Psychology of Leadership, Change, and Organizational Development*, 331-345.

Collins, C. S. (2015). Land-grant Extension: Defining public good and identifying pitfalls in evaluation. *Journal of Higher Education Outreach and Engagement*, *19*(2), 37-64.

Committee on Communication for Behavior Change in the 21st Century. (2002). *The National Academies of Sciences, Engineering, and Medicine.* Washington, DC.

Cook, N. E. (2016). *Aligning cultural responsiveness in evaluation and evaluation capacity building: A needs assessment with family support programs.* Unpublished Dissertation. Blacksburg, VA: Virginia Polytechnic Institute and State University.

Cousins, J. B., Goh, S. C., Elliott, C. J., & Bourgeois, I. (2014). Framing the capacity to do and use evaluation. *New Directions for Evaluation*, *2014*(141), 7-23.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Cummings, T. G., & Worley, C. G. (2014). *Organization development and change*. Boston, MA: Cengage Learning.

Daigle, J. J., Hrubes, D., & Ajzen, I. (2002). A comparative study of beliefs, attitudes, and values among hunters, wildlife viewers, and other outdoor recreationists. *Human Dimensions of Wildlife*, *7*(1), 1–19.

Davis, T. R. V., & Luthans, F. (1980). A social learning approach to organizational behavior. *The Academy of Management Review, 5*(2), 281-290.

Decker, D. J., & Yerka, B. L. (1990). Organizational philosophy for program evaluation. *Journal of Extension*, 28(2) Article 2FRM1. doi: http://www.joe.org/joe/1990summer/f1.php

DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.; Vol. 26). Los Angeles, CA: Sage Publications.

Diem, K. G. (2009). Preparing youth development professionals to be successful: How do the needs of Extension/4-H compare to those of other organizations? *Journal of Extension, 47*(1), Article 1RIB1. Online: htttp://www.joe.org/joe/2009february/rb1.php

Diem, K. G. (1997). Measuring impact of educational programs. *Rutgers Cooperative Extension fact sheet #869*. New Brunswick, NJ: Rutgers University.

Diem, K. G. (2003). Program development in a political world—it's all about impact! *Journal of Extension, 41*(1) Article 1FEA6. Online at: http://www.joe.org/joe/2003february/a6.shtml

Diem, K. G., Hino, J., Martin, D., & Meisenbach, T. (2011). Is Extension ready to adopt technology for delivering programs and reaching new audiences? *Journal of Extension, 49*(6), Article 6FEA1. Online: http://www.joe.org/joe/2011december/a1.php

Dillman, L. M. (2013). Evaluator skill-acquisition: Linking educational experiences to competencies. *American Journal of Evaluation, 34*(2), 270–285. doi: 10.1177/1098214012464512

Donald, I. J., Cooper, S. R., & Conchie, S. M. (2014). An extended theory of planned behaviour model of the psychological factors affecting commuters' transport mode use. *Journal of environmental psychology, 40*, 39-48.

Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications.* London, England: Routledge.

Donaldson, S. I., Gooler, L. E., & Scriven, M. (2002). Evaluation strategies for managing evaluation anxiety: Toward a psychology of program. *American Journal of Evaluation*, 23, 261–273.

D'Ostie-Racine, L., Dagenais, C., & Ridde, V. (2016). A qualitative case study of evaluation use in the context of a collaborative program evaluation strategy in Burkina Faso. *Health Research Policy and Systems*, *14*(1), 37.

Douglah, M., Boyd, H., & Gundermann, D. (2003). Nurturing the development of an evaluation culture in public educational agencies. In A*nnual Conference of the American Evaluation Association, Reno, NV*.

Dozois, E., Langlois, M., & Blancher-Cohen, N. (2010). *DE 201: A practitioner's guide to developmental evaluation.* Montreal, Canada: The J. W. McConnell Family Foundation.

Duttweiler, M. W. (2008). The value of evaluation in Cooperative Extension. In M. T. Braverman, M. Engle, M. E. Arnold, & R. A. Rennekamp (Eds.), *Program evaluation in a complex organizational system: Lessons from Cooperative Extension. New Directions for Evaluation, 120,* 87–100. doi: https://doi.org/10.1002/ev.278

Earl, S., Carden, F., & Smutylo, T. (2001). *Outcome mapping: Building learning and reflection into development programs*. IDRC, Ottawa, ON, CA.

European Evaluation Society. (2007). *The importance of a methodologically diverse approach to impact evaluation.* Online at: http://www.europeanevaluation.org/news?newsId_1969406

eXtension (eXtension Foundation**).** (2017) Impact evaluation course. Impact Collaborative. USDA National Institute of Food and Agriculture (FIFA). Online: https://www.extension.org/impact-collaborative/

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology*, *23*, 75-109.

Fazio, R. H. (2013). The attitude-to-behavior process: Implications. *Advertising Exposure, Memory and Choice*, 119-129.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*(6), 1013.

Fazio, R. H., & Olson, M. A. (2014). The MODe Model. *Dual-process theories of the social mind*, 155.

Fazio, R. H., Powell, M. C., & Williams, C. J. (1989). The role of attitude accessibility in the attitude-to-behavior process. *Journal of Consumer Research*, *16*(3), 280-288.

Fetterman, D., Kaftarian, S., & Wandersman, A. (Eds.). (1996). *Empowerment evaluation: Knowledge and tools for self-assessment and accountability*. Thousand Oaks, CA: Sage Publications.

Fierro, L. A., & Christie, C. A. (2016). Evaluator and program manager perceptions of evaluation capacity and evaluation practice. *American Journal of Evaluation*, 1098214016667581.

Fishbein, M., & Ajzen, I. (2010). Predicting and changing behavior: The reasoned action approach. New York, NY: Psychology Press.

Frank, A.W., 2010. *Letting stories breathe: a socio-narratology*. Chicago, IL: The University of Chicago Press. Doi: 10.7208/chicago/9780226260143.001.0001.

Franz, N. (2009). A holistic model of engaged scholarship: Telling the story across higher education's missions. *Journal of Higher Education Outreach and Engagement, 13*(4), 31–43.

Franz, N. (2011). Advancing the public value movement: Sustaining Extension during tough times. *Journal of Extension, 49*(2). Online: http://www.joe.org/joe/2011april/comm2.php

Franz, N. K., Arnold, M., & Baughman, S. (2014). The role of evaluation in determining the public value of Extension. *Journal of Extension, 52*(4).

French, C., & Morse, G. (2015). Extension stakeholder engagement: An exploration of two cases exemplifying 21st Century adaptions. *Journal of Human Sciences and Extension Volume, 3*(2), 108-131.

Frey, B. B. (Ed.). (2018). *The SAGE encyclopedia of educational research, measurement, and evaluation*. SAGE Publications.

Forss, K., Kruse, S. E., Taut, S., & Tenden, E. (2006). Chasing a ghost? An essay on participatory evaluation and capacity development. *Evaluation*, *12*(1), 128-144.

Fouts, H. (2004). *Organizational climate of North Carolina Cooperative Extension*, PhD dissertation. Raleigh, NC: North Carolina State University. ProQuest Dissertations Publishing, # 3154202. ISBN: 9780496146994

Fowler, F. J., Jr. (2014). *Survey research methods* (5th Ed.). Thousand Oaks, CA: Sage Publications, Inc.

Franco, L. M., Bennett, S., & Kanfer, R. (2002). Health sector reform and public sector health worker motivation: a conceptual framework. *Social Science & Medicine, 54*(8), 1255-1266.

Franz, N. K. (2013). Improving Extension programs: Putting public value stories and statements to work. *Journal of Extension, 51*(3), 3TOT1-None.

Franz, N., & Townson, L. (2008). The nature of complex organizations: The case of Cooperative Extension. In M. T. Braverman, M. Engle, M. E. Arnold, & R. A. Rennekamp (Eds.), Program evaluation in a complex organizational system: Lessons from Cooperative Extension (pp. 5–14). *New Directions for Evaluation, 2008*(120). San Francisco, CA: Jossey-Bass. doi:10.1002/ev.272

Fullan, M. (1993). *Change forces: Probing the depths of educational reform* (Vol. 10). Psychology Press.

Funnell, S. C. (2000). Developing and using a program theory matrix for program evaluation and performance monitoring. *New Directions for Evaluation, 2000*(87), 91-101.

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models* (Vol. 31). John Wiley & Sons.

Gargani, J. (2013). What can practitioners learn from theorists' logic models? *Evaluation and Program Planning, 38*, 81-88.

Gertler, P., Martinez, S., Premand, P., Rawlings, L., & Vermeersch, C. (2011). *Impact evaluation in practice*. Washington, CD: The International Bank for Reconstruction and Development / The World Bank. DOI: 10.1596/978-0-8213-8541-8

Ghimire, N., Koundinya, V., & Holz-Clause, M. (2014). Government Run vs. University Managed Agricultural Extension: A Review of Nepal, India and the United States. *Asian Journal of Agricultural Extension, Economics & Sociology 3 (5),* 461-472.

Ghimire, N. R., & Martin, R. A. (2011). A professional competency development model: Implications for Extension educators. *Journal of International Agricultural and Extension Education, 18*(2), 5-17.

Ghimire, N. R., & Martin, R. A. (2012). Learning systems competencies: Their importance to Extension educators. *International Journal of Learning*, *18*(3).

Ghimire, N. R., & Martin, R. A. (2013). Does evaluation competence of Extension educators differ by their program area of responsibility? *Journal of Extension*, *51*(6), n6.

Ghimire & Martin (2015). To what degree do extension educators self-assess their needs for building evaluation capacity? A case study of the north central region of the United States. ResearchGate. doi: 311931658

Grace, J. & Bollen, K. (2005). Interpreting the results from multiple regression and structural equation models. *The Bulletin of the Ecological Society of America*, *86*(4), 283-295.

Greaves, M., Zibarras, L. D., & Stride, C. (2013). Using the theory of planned behavior to explore environmental behavioral intentions in the workplace. *Journal of Environmental Psychology, 34,* 109–120. doi: http://doi.org/10.1016/j.jenvp.2013.02.003

Green, K. E., & Frantom, C. G. (2002, November). Survey development and validation with the Rasch model. In *International Conference on Questionnaire Development, Evaluation, and Testing,* Charleston, SC.

Griffin, L. M. (2015). *An evaluation of the military extension internship program*. Unpublished thesis. Purdue University.

Guba, E. G. and Lincoln, Y. (2001). Guidelines and checklist for constructivist evaluation. Kalamazoo, MI: Evaluation Centre. Online at: www.wmich.edu/evalctr/checklists.

Guion, L., Boyd, H., & Rennekamp, R. (2007). An exploratory profile of Extension evaluation professionals. *Journal of Extension, 5(4)*. https://www.joe.org/joe/2007august/a5.php

Han, H. (2015). Travelers' pro-environmental behavior in a green lodging context: Converging value-belief-norm theory and the theory of planned behavior. *Tourism Management, 47*, 164-177.

Han, H., & Kim, Y. (2010). An investigation of green hotel customers' decision formation: Developing an extended model of the theory of planned behavior. *International Journal of Hospitality Management, 29*(4), 659-668.

Hankins, M., French, D., & Horne, R. (2000). Statistical guidelines for studies of the theory of reasoned action and the theory of planned behaviour. *Psychology & Health, 15*(2), 151-161

Harder, A., Place, N. T., & Scheer, S. D. (2010). Towards a competency-based Extension education curriculum: A Delphi study. *Journal of Agricultural Education, 51*(3), 44–52. doi:10.5032/jae.2010.03044

Harder, A., Gouldthorpe, J., & Goodwin, J. (2015a). Exploring organizational factors related to Extension employee burnout. *Journal of Extension, 53*(2), 2FEA2.

Harder, A., Scheer, S., & Place, N. (2015b). Improving the future of US Cooperative Extension: A Delphi study. *Journal of Extension Systems, 27*(2).

Hatry, H. (1999). Introduction. *Public Administration Review, 59*(2), 101-104.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research, 15*(9), 1277-1288.

Henry, G.T., & Mark, M.M. (2003a). Beyond use: Understanding evaluation's influence on attitudes and actions. *The American Journal of Evaluation*, *24*(3), 293-314. doi: https://doi.org/10.1016/S1098-2140(03)00056-0

Henry, G.T., & Mark, M.M. (2003b). Toward an agenda for research on evaluation. *New Directions for Evaluation, 97,* 69-80. doi: 10.1002/ev.77

Higgins, C. A., Duxbury, L. E., & Lyons, S. T. (2010). Coping with overload and stress: Men and women in dual-earner families. *Journal of Marriage and Family, 72*(4), 847-859.

Holbert, R. L., & Stephenson, M. T. (2002). Structural equation modeling in the communication sciences, 1995–2000. *Human Communication Research*, *28*(4), 531-551.

Hodge, B., Wright, B., & Bennett, P. (2016). Does academic training change intentions? Drawing upon the theory of planned behaviour to improve academic performance. *International Journal of Training Research*, *14*(1), 1-14. doi: 10.1080/14480220.2016.1259004

Howes, M. (1992). Linking paradigms and practice: Key issues in the appraisal, monitoring and evaluation of British NGO projects. *Journal of International Development*, *4*(4), 375-396.

Huetteman, J. (2017). *Outcomes & impacts: Giving programs focus for evaluation*. Unpublished PowerPoint presentation. Purdue Extension Service.

Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorizing in implicit attitude research: Propositional and behavioral alternatives. *The Psychological Record*, *61*(3), 465.

Israel, G. D. (1992). *Sampling the evidence of extension program impact*. University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation, 30*(3), 377-410. doi: 10.1177/1098214009341660

Jones, N., Jones, H., Steer, L., & Datta, A. (2009). *Improving impact evaluation production and use*. London: Overseas Development Institute.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 294-316). Newbury Park, CA: Sage.

Kalambokidis, L. (2004). Identifying the public value in Extension programs. *Journal of Extension, 42*(2), Article 2FEA1. Online: http://www.joe.org/joe/2004april/a1.php

Kalambokidis, L. (2011). Spreading the word about Extension's public value. *Journal of Extension, 49*(2), 2FEA1. Online: http://www.joe.org/joe/2011april/a1.php

Kalambokidis, L. (2017). Building Extension's public value. Conference presentation. Fargo, ND: NDSU Extension REC and Annual Conference.

Khandker, S., B. Koolwal, G., & Samad, H. (2009). *Handbook on impact evaluation: quantitative methods and practices*. The World Bank.

Keeling, R. P., Wall, A. F., Underhile, R., & Dungy, G. J. (2008). Assessment reconsidered: Institutional effectiveness for student success. *International Center for Student Success and Institutional Accountability, 78.*

Kelsey, K. D. (2008). Do workshops work for building evaluation capacity among Cooperative Extension Service faculty? *Journal of Extension, 46*(6). doi: http://www.joe.org/joe/2008october/2008december/pdf/JOE_v46_6rib4.pdf

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural equation modeling, 10*(3), 333-351.

King, J. A. (2007). Developing evaluation capacity through process use. In J. B. Cousins (Ed.), Process use in theory, research, and practice (pp. 45–59). *New Directions for Evaluation*, *2007*(116). doi:10.1002/ev.242

King, J. A., & Stevahn, L. (2013). Essential competencies for program evaluators. In *Workshop presented at the Summer Institute of the American Evaluation Association, Atlanta, GA*.

King, J. A., & Stevahn, L. (2015). Competencies for program evaluators in light of adaptive action: What? So what? Now what? *New Directions for Evaluation*, *2015*(145), 21-37.

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of American Society of Training Directors, 11*, 1-13.

Kirkpatrick, D. (1996, January). Great ideas revisited. *Training & Development*,*50*(1), 54-59.

Kirkpatrick, D. L. (2009). *Implementing the Four Levels: A Practical Guide for Effective Evaluation of Training Programs: Easyread Super Large 24pt Edition*. ReadHowYouWant.com.

Kirkpatrick, D. (2010). 50 years of evaluation. *Training and Development, 64*(1), 14.

Kirkpatrick, D., & Kirkpatrick, J. (2005). *Transferring learning to behavior: Using the four levels to improve performance*. Berrett-Koehler Publishers.

Kirkpatrick, D., & Kirkpatrick, J. (2006). *Evaluating training programs: The four levels*. Berrett-Koehler Publishers.

Kitinoja, L. (1989). Effects of perceived involvement of stakeholders on readiness for utilization of program evaluation results by county administrators in the U.S. Cooperative Extension Service. Paper presented at the Annual Meeting of the American Evaluation Association, San Francisco, CA: October 19–21.

Kline, R. B. (2016). *Principles and practices of structural equation modeling* (4th ed.). New York: Guilford.

Knowles, M. S. (1990). *The adult learner: A neglected species* (4th ed.). Houston: Gulf.

Knowles, M.S., Holton, E. F., & Swanson, R. A. (1998). *The adult learner* (5th ed.). Houston: Gulf.

Kollock, D. H., Flage, L., Chazdon, S., Paine, N., & Higgins, L. (2012). Ripple effect mapping: A "radiant" way to capture program impacts. *Journal of Extension*, *50*(5), 5TOT6.

Kluchinski, D. (2014). Evaluation behaviors, skills and needs of Cooperative Extension agricultural and resource management field faculty and staff in New Jersey. *Journal of the NACAA*, *7*(1).

Kuchinke, K. P. (1995). Managing learning for performance. *Human Resource Development Quarterly*, *6*(3), 307-316.

Labin, S. N. (2014). Developing common measures in evaluation capacity building: An iterative science and practice process. *American Journal of Evaluation*, *35*(1), 107-115.

Labin, S. N., Duffy, J. L., Meyers, D. C., Wandersman, A., & Lesesne, C. A. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation*, *33*(3), 307-338.

Lam, T., & Hsu, C. H. (2006). Predicting behavioral intention of choosing a travel destination. *Tourism Management*, *27*(4), 589-599.

Lambur, M. T. (2008). Organizational structures that support internal program evaluation. In M. T. Braverman, M. Engle, M. E. Arnold, & R. A. Rennekamp (Eds.), Program evaluation in a complex organizational system: Lessons from Cooperative Extension (pp. 41–54).

New Directions for Evaluation, 2008(120). San Francisco, CA: Jossey-Bass. doi:10.1002/ev.275

Lamm, A, J, (2011), Effect of organizational context on extension evaluation behaviors. Unpublished dissertation. University of Florida. Gainesville: FL.

Lamm, K. W., Carter, H. S., & Lamm, A. J. (2016). Evaluating Extension based leadership development programs in the southern United States. *Journal of Agricultural Education*, *57*(1), 121-136.

Lamm, A. J., & Israel, G. D. (2011). An organizational approach to understanding evaluation in Extension. *Journal of Agricultural Education*, *52*(4), 136-149.

Lamm, A. J., Israel, G. D., & Diehl, D. (2013). A national perspective on the current evaluation activities in Extension. *Journal of Extension*, *51*(1), 1-6.

Lamm, A. J., Israel, G. D., & Harder, A. (2011). Getting to the bottom line: How using evaluation results to enhance Extension programs can lead to greater levels of accountability. *Journal of Agricultural Education*, *52*(4), 44-55.

Larson. M. (2016). Perceptions of evaluator competencies in public health. Dissertation. University of Minnesota, ProQuest Dissertations Publishing, 2016. 10141994.

Leeuw, F. L., & Vaessen, J. (2009). *Impact evaluations and development: NONIE guidance on impact evaluation*. Network of networks on impact evaluation.

Lekies, K. S., & Bennett, A. M. (2011). The evaluation attitudes and practices of 4-H educators. *Journal of Extension*, *49*(1), 1RIB2.

Lin, J., Chan, H. C., & Wei, K. K. (2006). Understanding competing application usage with the theory of planned behavior. *Journal of the Association for Information Science and Technology*, *57*(10), 1338-1349.

Luzar, E. J., & Diagne, A. (1999). Participation in the next generation of agriculture conservation programs: The role of environmental attitudes. *The Journal of Socio-Economics, 28*(3), 335-349.

MacQueen, K. M., McLellan-Lemal, E., Bartholow, K., & Milstein, B. (2008). Team-based codebook development: Structure, process, and agreement. In G. Guest & K. M. MacQueen (Eds.), *Handbook for team-based qualitative research* (pp. 119–35). Lanham, MD: AltaMira Press.

Madden, T. J., Ellen, P. S., & Ajzen, I. (1992). A comparison of the theory of planned behavior and the theory of reasoned action. *Personality and Social Psychology Bulletin*, *18*(1), 3-9.

Mark, M. M., Donaldson, S. I., & Campbell, B. (Eds.). (2011). *Social Psychology and Evaluation*. New York: The Guilford Press.

Marks, S. R. (1977). Multiple roles and role strain: Some notes on human energy, time and commitment. American Sociological Review, 42, 921– 936.

Marks, S.R., & McDermid, S.M. (1996). Multiple roles and the self: A theory of role balance. Journal of Marriage and the Family, 58(2), 417-433.

Marshall, J. P., Higginbotham, B. J., Harris, V. W., & Lee, T. R. (2007). Assessing program outcomes: Rationale and benefits of posttest-then-retrospective-pretest designs. *Journal of Youth Development*, *2*(1), 118-123.

McClure, M. M., Fuhrman, N. E., & Morgan, A. C. (2012). Program Evaluation Competencies of Extension Professionals: Implications for Continuing Professional Development. *Journal of Agricultural Education, 53*(4), 85-97. doi: 10.5032/jae.2012.04085

McEachan, C., Conner, M., Taylor, N., & Lawton, R. (2011). Prospective prediction of health-related behavior with the TPB: a meta-analysis. *Health Psychology Review 5*(2), 97–144.

McGaan, Lee. "Persuasion Theory Intro [101]." *Introduction to Persuasion*. Monmouth University, 2009. Web. 21 Jan. 2015.

McKim, B. R. (2013, December). Assessing knowledge, performance, and consequence competence with the Borich Needs Assessment Model. Texas A&M University.

McKim, B. R., & Saucier, P. R. (2011). An Excel-based mean weighted discrepancy score calculator. *The Journal of Extension*, *49*(2). Retrieved from http://www.joe.org/joe/2011april/pdf/JOE_v49_2tt8.pdf

McLeroy, K.R., Bibeau, D., Steckler, A., & Glanz, K. (1988). An ecological perspective on health promotion programs. Health Education Quarterly, 15, 351-377.

Meng, C. L., Othman, J., D'Silva, J. L., & Omar, Z. (2014). Ethical decision making in academic dishonesty with application of modified theory of planned behavior: A review. *International Education Studies*, *7*(3), 126.

Miller, R. L. (2013). Logic models: A useful way to study theories of evaluation practice? *Evaluation and Program Planning, 38*, 77–80

Minarovic, R. E., & Mueller, J. P. (2000). North Carolina Cooperative Extension Service professionals' attitudes toward sustainable agriculture. Journal of Extension [Online], 38(1), Article 1FEA1. Online: http://www.joe.org/joe/2000february/a1.php

Mohr, L. B. (1995). *Impact analysis for program evaluation (2nd Ed.)*. Sage.

Montaño, D. E., & Kasprzyk, D. (2015). Theory of reasoned action, theory of planned behavior, and the integrated behavioral model. *Health behavior: Theory, research and practice*, 95-124.

Morford, S., Kozak, R., Suvedi, M., & Innes, J. (2006). Factors affecting program evaluation behaviours of natural resource Extension practitioners—motivation and capacity building. J*ournal of Extension, 44*(3), Article 3FEA7.

National Professional Development Task Force. (2004). *New foundations for the 4-H youth development profession.* Washington, DC: National 4-H Council.

Network of Networks on Impact Evaluation. (2007). *Impact evaluation guidance*, working draft, subgroup 2 (SG2). Online at: http://www.worldbank.org/ieg/nonie/index.html

Nichols, A., Blake, S. M., Chazdon, S., & Radhakrishna, R. (2015). From farm results demonstrations to multistate impact designs: Cooperative Extension navigates its way through evaluation pathways. *Journal of Human Sciences and Extension Volume, 3*(2).

Oh, H., & Hsu, C. H. (2001). Volitional degrees of gambling behaviors. *Annals of Tourism Research*, *28*(3), 618-637.

Olson, K. D., Skuza, J. A., & Blinn, C. R. (2007). Extension educators' views of scholarship and performance evaluation criteria. *Journal of Extension*.

O'Neill, B. (1998). Money talks: Documenting the economic impact of Extension personal finance programs. Journal of Extension [On-line], 36(5) Article 5FEA2. Available at: http://www.joe.org/joe/1998october/a2.php

Orbeil, S., Hodgkins, S., & Sheeran, P. (1997). Implementation intentions and the theory of planned behavior. *Personality and Social Psychology Bulletin*, *23*(9), 945–954. doi: http://doi.org/10.1177/0146167297239004

O'Sullivan, R. G. (2012). Collaborative evaluation within a framework of stakeholder-oriented evaluation approaches. *Evaluation and Program Planning*, *35*(4), 518–522. Doi: http://doi.org/10.1016/j.evalprogplan.2011.12.005

Patton, M. Q. (1982). *Practical evaluation*, Newbury Park, CA: Sage Publications.

Patton, M. Q. (2003). Utilization-focused evaluation. In T. Kellaghan, & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 223–244). Netherlands: Kluwer.

Patton, M. Q. (2007). Process use as a usefulism. *New Directions for Evaluation*, *2007*(116), 99-112.

Patton, M. Q. (2008). sup wit eval ext? In M. T. Braverman, M. Engle, M. E. Arnold, & R. A. Rennekamp (Eds.), *Program evaluation in a complex organizational system: Lessons from Cooperative Extension. New Directions for Evaluation, 120,* 101–115.

Patton, M. Q. (2008). *Utilization-focused evaluation*. 4th Ed. Sage.

Parkinson, S. (2009). Power and perceptions in participatory monitoring and evaluation. *Evaluation and Program Planning*, *32*(3), 229–237. http://doi.org/10.1016/j.evalprogplan.2009.03.003

Patton, M. Q. (2010). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. Guilford Press.

Patton, M. Q. (2012). *Essentials of utilization-focused evaluation*. Sage.

Picciotto, R. (2014) Is impact evaluation? European Journal of Development Research 26 (1): 31–38.

Prochaska, J., DiClemente, C., & Norcross, J. (1992). In search of how people change. *American Psychologist*, 47, 1102-1114.

Putnam, R. (1993). The prosperous community: Social capital and public life. *The American Prospect, 13*, 35–42.

Peterson, D. J., & Downey, L. H. (2014). National 4-H healthy living program environmental scan and evaluability assessment report. NIFA: USDA. Washington, D.C.

Preskill, H. (2014). Now for the hard stuff: Next steps in ECB research and practice. *American Journal of Evaluation, 35*(1), 116-119. doi: 1098214013499439.

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, *29*(4), 443-459.

Preskill, H., Martinez-Papponi, B., & Torres, R. T. (2001, February). Organizational Readiness for Learning and Evaluation. *In Academy of Human Resource Development 2001 Conference Proceedings.*

Preskill, H., & Russ-Eft, D. (2015). *Building evaluation capacity: Activities for teaching and training*. Sage Publications.

Preskill, H., & Torres, R. T. (2000). The readiness for organizational learning and evaluation instrument (ROLE). *Evaluation in Organizations*, 421-434.

Preskill, H., Torres, R. T., & Martinez-Papponi, B. (1999). *Assessing an organization's readiness for learning from evaluative inquiry*. Paper presented at the American Evaluation Association annual conference. Orlando, FL, November.

Prochaska, J., DiClemente, C., & Norcross, J. (1992). In search of how people change. *American Psychologist*, 47, 1102-1114.

Putnam, R. (1993). The prosperous community: Social capital and public life. *The American Prospect, 13*, 35–42.

Rennekamp, R. A., & Arnold, M. E. (2009). What progress, program evaluation? Reflections on a quarter-century of Extension evaluation practice. *Journal of Extension, 47*(3), Article 3COM1. Online: http://www.joe.org/joe/2009june/comm1.php

Rennekamp, R. A., & Engle, M. (2008). A case study in organizational change: Evaluation in Cooperative Extension. In M. T. Braverman, M. Engle, M. E. Arnold, & R. A.

Rennekamp (Eds.), Program evaluation in a complex organizational system: Lessons from Cooperative Extension (pp. 15–26). *New Directions for Evaluation, 2008*(120). San Francisco, CA: Jossey-Bass. doi:10.1002/ev.273

Riessman, C., 2008. *Narrative methods for the human sciences*. Thousand Oaks, CA: Sage.

Roche, C. J. (1999). *Impact assessment for development agencies: Learning to value change*. ISBN 9780855984182. Oxfam. Oxford: UK.

Rockwell, S. (2002). Targeting Outcomes of Programs (TOP): A hierarchy for targeting outcomes and evaluating their achievement. PowerPoint Presentation. *Faculty Publications: Agricultural Leadership, Education & Communication Department*. 49. Lincoln, NE: University of Nebraska.

Rockwell, K., & Bennett, C. (2004). Targeting outcomes of programs: A hierarchy for targeting outcomes and evaluating their achievement.

Rodgers, M. S., Hillaker, B. D., Haas, B. E., & Peters, C. (2012). Taxonomy for Assessing Evaluation Competencies in Extension. *Journal of Extension*, *50*(4).

Rog, D. J. (2012). When background becomes foreground: Toward context-sensitive evaluation practice. In D. J. Rog, J. L. Fitzpatrick, & R. F. Conner (Eds.), *Context: A frame-work for its influence on evaluation practice. New Directions for Evaluation*, *135*, 25–40.

Rogers, A. F., & Gullickson, A. (2018). Evaluation Champions: A Literature Review. *Journal of MultiDisciplinary Evaluation*, *14*(30), 46-63.

Rogers, A. F., Kelly, L. M., & McCoy, A. (2019). Pathways to becoming an internal evaluator: Perspectives from the Australian non-government sector. *Evaluation and Program Planning*. In press. doi: https://doi.org/10.1016/j.evalprogplan.2019.01.007

Rogers, Everett M. (2010). *Diffusion of innovations* (3d Ed.). New York: Free Press.

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Thousand Oaks, CA: Sage.

Sanbonmatsu, D. M., & Fazio, R. H. (1990). The role of attitudes in memory-based decision making. *Journal of Personality and Social Psychology, 59*(4), 614.

Sandmann, L. R. (2008). Conceptualization of the scholarship of engagement in higher education: A strategic review, 1996–2006. *Journal of Higher Education Outreach and Engagement*, *12*(1), 91-104.

Schein, E. H. (1990). *Organizational culture* (Vol. 45, No. 2, p. 109). American Psychological Association.

Scheirer, M. A. (2005). Is sustainability possible? A review and commentary on empirical studies of program sustainability. *American Journal of Evaluation, 26*(3), 320–347. doi:10.1177/1098214005278752

Schlüter, M., Baeza, A., Dressler, G., Frank, K., Groeneveld, J., Jager, W, Janssen, M.A., McAllister, R.R., Müller, B., Orach, K., & Schwarz, N. (2017). A framework for mapping and comparing behavioural theories in models of social-ecological systems. *Ecological Economics*, *131*, 21-35.

Schuette, R. A., & Fazio, R. H. (1995). Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model. *Personality and Social Psychology Bulletin*, *21*(7), 704-710.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, *26*(3-4), 207-231.

Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. In A. Wigfield & J. S. Eccles (Eds.), *Development of Achievement Motivation*, (pp. 15-31). San Diego, CA: Academic Press.

Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition, 25*(5), 638-656.

Seevers, B., Graham, D., & Conklin, N. (2007). *Education through Cooperative Extension* (2nd ed.). Columbus: The Ohio State University.

Senge, P. M. (1990). Senge, P. (1990). *The fifth discipline: The art & practice of the learning organization.* New York: Doubleday.

Shaw, J., & Campbell, R. (2014). The "process" of process use: Methods for longitudinal assessment in a multisite evaluation. *American Journal of Evaluation*, *35*(2), 250-260.

Shaw, I., Shaw, I. G. R., Greene, J. C., & Mark, M. M. (Eds.). (2006). *The Sage handbook of evaluation*. Sage.

Silliman, B., Crinion, P., & Archibald, T. (2016a). Evaluation Champions: What They Do, Why They Do It, and Why It Matters to Organizations. *Journal of Human Sciences and Extension*, *4*(3), 1-21.

Silliman, B., Crinion, P., & Archibald, T. (2016b). Evaluation champions: What they need and where they fit in organizational learning. *Journal of Human Sciences and Extension Volume*, *4*(3), 22-44.

Smith, A. Y., (2015). *Attitude, subjective norm, and perceived behavioral control as indicators for nurse educators' intention to use critical thinking teaching strategies: A structural equation model analysis.* Dissertations. Paper 1576. Andrews University. Berrien Springs, MI.

Sparkes, A.C. and Smith, B., 2012. Narrative analysis as an embodied engagement with the lives of others. *In*: J. Holstein and J. Gubrium, eds. *Varieties of analysis.* London: Sage, 53–74. doi: 10.4135/9781506335117.

Stame, N., & Presti, V. L. (2017). Positive thinking and learning from evaluation. In *Success in Evaluation* (pp. 19-45). Routledge.

StataCorp (2017). *Stata structural equation modeling reference manual*. College Station, TX: Stata Press.

Stone B., & Bieber, S. (1997). Competencies: A new language for our work. *Journal of Extension*, *35*(1).

Stone, B., & Coppernoll, S. (2004). You, extension and success: A competency-based professional development system. *Journal of Extension*, *42*(2).

Suvedi, M. (2011). Evaluation of agricultural extension and advisory services: A MEAS manual. Michigan State University and the USAID. Online: www.meas-extension.org/measoffers/training/evaluatingextensionprograms

Suvedi, M., & Kaplowitz, M. D. (2016). *What every Extension worker should know: Core competency handbook.* Michigan State University, Department of Community Sustainability.

Tabachnick, B. G., Fidell, L. S., & Osterlind, S. J. (2001). Using multivariate statistics. Online (501802917): http://tocs.ulb.tu-darmstadt.de/135813948.pdf

Taut, S. (2007). Studying self-evaluation capacity building in a large international development organization. *American Journal of Evaluation, 28*(1), 45-59.

Taut, S., & Brauns, D. (2003). Resistance to evaluation: A psychological perspective. *Evaluation, 9*(3), 247-264.

Taylor-Powell, E., Jones, L., & Henert, E. (2003). *Enhancing program performance with logic models.* Online course: http://www.uwex.edu/ces/lmcourse/ Madison, WI: University of Wisconsin-Extension.

Taylor-Powell, E., Steele, S., & Douglah, M. (1996). *Planning a program evaluation.* Madison, WI: University of Wisconsin Cooperative Extension.

Themessl-Huber, M., Harrow, A., & Laske, S. (2005). Evaluator roles. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 147–149). Thousand Oaks, CA: Sage.

Thomas, V. & Parsons, B. (2017). Culturally responsive evaluation meets systems-oriented evaluation. *American Journal of Evaluation, 38*(1), 7-28.

Trevino, L. K. (1986). Ethical decision making in organizations: A person-situation interactionist model. *Academy of Management Review*, *11*(3), 601-617.

Trochim, W., Urban, J. B., Hargraves, M., Hebbard, C., Buckley, J., Archibald, T., & Burgermaster, M. (2012). *The guide to the systems evaluation protocol.* Ithaca, NY: Cornell Digital Print Services.

United Way of America. (1995). *Current United Way approaches to measuring program outcomes and community change.* Alexandria, VA: United Way of America.

United Way of America. (1996). *Measuring program outcomes: A practical approach*. Arlington, VA: United Way of America.

University of Wisconsin-Extension (UWEX). (1995). *Program logic model.* Madison, WI: UWEX.

Urban, J. B., Burgermaster, M., Archibald, T., & Byrne, A. (2015). Relationships between quantitative measures of evaluation plan and program model quality and a qualitative measure of participant perceptions of an evaluation capacity building approach. *Journal of Mixed Methods Research, 9*(2), 154-177.

USAID Center for Development Information and Evaluation. (2017). U.S. Agency for International Development (USAID). Online: https://www.usaid.gov/evaluation

USDA NIFA. (2017). *Extension*. Online: http://www.csrees.usda.gov/qlinks/extension.html

Van der Horst, M. (2016). Role theory. *Sociology-Oxford Bibliographies*. doi: https://doi.org/10.1093/OBO/9780199756384-0175

Van Twillert, A. 2016. Factors and Beliefs that Influence Higher Education Teachers' Intention to Integrate Web 2.0 Technologies into Their Constructivist Pedagogy: A Study with the Reasoned Action Approach, Master's Thesis. Open University, The Netherlands.

Var, I. (1998). Multivariate data analysis. *vectors*, *8*(2), 125-136.

Vengrin, C. A. (2016). *Examining the evaluation capacity, evaluation behaviors, and the culture of evaluation in Cooperative Extension.* Unpublished Dissertation. Blacksburg, VA: Virginia Polytechnic Institute and State University.

Vengrin, C., Westfall-Rudd, D., Archibald, T., Rudd, R., & Singh, K. (2018). Factors affecting evaluation culture within a non-formal educational organization. *Evaluation and program planning, 69,* 75-81.

Volkov, B. B. (2011). Beyond being an evaluator: The multiplicity of roles of the internal evaluator. *New Directions for Evaluation*, *2011*(132), 25-42.

Vroom, V. H. (1964). *Work and motivation.* New York: John Wiley & Sons.

Waddell, D. L. (1991). Differentiating impact evaluation from evaluation research: One perspective of implications for continuing nursing education. *The Journal of Continuing Education in Nursing, 22*(6), 254-258.

Wandersman, A., Imm, P., Chinman, M., & Kaftarian, S. (2000). Getting to outcomes: A results-based approach to accountability. *Evaluation and Program Planning, 23,* 389-395.

Warner, P. D., & Christenson, J. A. (1984). *The Cooperative Extension Service: a national assessment*. USDA.

Weiss, C.H. (1998). *Evaluation.* Upper Saddle River, NJ: Prentice-Hall, Inc.

Wells-Marshall, J. A. (2012). *The Extension program development model: A look at commitment and engagement*. Doctoral dissertation. Auburn University.

West, J., & Bogers, M. (2014). Leveraging external sources of innovation: a review of research on open innovation. *Journal of Product Innovation Management, 31*(4), 814-831.

Wholey, J. (1987). Evaluability assessment: Developing program theory. In L. Bickman, (Ed.), *Using Program Theory in Evaluation, New Directions for Program Evaluation*. 33, 77-92. San Francisco, CA: Jossey-Bass Publishers.

Wilcox, Y., & King, J.A. (2013). A professional grounding and history of the development and formal use of evaluator competencies. *The Canadian Journal of Program Evaluation, 28*(3), 1-28).

Winfrey, E. C. (1999). Kirkpatrick's four levels of evaluation. In B. Hoffman (Ed.), *Encyclopedia of educational technology*. Online: http://coe.sdsu.edu/eet/Articles/k4levels/start.htm

Wippersberg, J. (2017). Evaluation Research. *The International Encyclopedia of Communication Research Methods*, 1-10.

W. K. Kellogg Foundation. (2004). *Logic model development guide*. Battle Creek, MI: W. K. Kellogg Foundation.

W. K. Kellogg Foundation. (2010). *W.K. Kellogg Foundation evaluation handbook: Philosophy and expectations.* Battle Creek, MI: W. K. Kellogg Foundation.

Workman, J. D. (2010). *Exploring the level of evidence of impact collected through program evaluation by Cooperative Extension as published in the Journal of Extension.* Doctoral dissertation, The Ohio State University. ProQuest Dissertations Publishing, 2010. 3438221.

Workman, J. D., & Scheer, S. D. (2012). Evidence of impact: Examination of evaluation studies published in the Journal of Extension. *Journal of Extension*, *50*(2).

World Bank. (1997). *The international bank for reconstruction and development.* Washington, DC: The World Bank.

APPENDIX A. DATA COLLECTION INSTRUMENT

NDSU NORTH DAKOTA STATE UNIVERSITY

Q1.1.

Study Introduction & Consent Form

Title of Research Study:
Program Impact Evaluation in Cooperative Extension: Competencies, Culture and Behaviors

This study is being conducted by: Jan Flack, Doctoral Student, and Dr. Claudette Peterson, Adviser.

Why I am being asked to take part in this research study? You are invited to participate in this research because you are a professional Extension Educator/Agent, Specialist, or Administrator within the North Central Region of the United States **who supports and/or engages in doing impact evaluation of educational programs**.

What is the reason for doing the study?
The study explores the factors that influence Extension professionals' efforts to do impact evaluation of educational programs. Results may indicate individual and organizational strengths and challenges for how to support impact evaluation.

What I am being asked to do?
You are invited to complete this online survey. Questions explore your perceptions of your capacity to do impact evaluation and perceptions of the support & assistance for impact evaluation available in your work context. Answers will be kept confidential.

How long will the survey take?
Survey completion will take 20 to 25 minutes. Your answers will be saved and, if you cannot complete the survey the first time, you may access it again using the same link provided initially to complete it.

What are the benefits of participating? You will have an opportunity to reflect on your own impact evaluation experiences and your responses may inform future Extension impact evaluation efforts.

What are the risks and discomforts, if any? Completion of the survey will involve no undue risk or discomfort. You may choose not to answer any question, for any reason. Your answers to all questions will be kept completely confidential and no personal identifying questions will be asked.

Do you have to take part in the study? It is your choice whether or not to participate.

What if I have questions?
This study has been reviewed by the NDSU Institutional Review Board (IRB). If you have any questions about your rights as a human research participant, or if you wish to report a concern, please contact the NDSU Human Research Protection Program at:

- Telephone: 701.231.8995 · Toll Free: 855.800.6717 · Email: ndsu.irb@ndsu.edu
- Mail: NDSU HRPP, 1735 NDSU Research Park Dr., NDSU Dept. 4000, PO Box 6050, Fargo, ND 58108-6050

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/research/irb.

Documentation of Informed Consent:
You are freely making a decision whether to be in this research study. Continuing with this survey means that:

1. You have read and understood this consent form
2. You have decided to be in the study.

>> By clicking on the **Arrow** below you agree to participate in this study.

~ ~ ~ ~ ~ ~ ~ ~ ~ ~

Please note: This survey will be best displayed on a laptop or desktop computer.
Some features may be less compatible for use on a mobile device.

→

Q2.1.
Study Focus: Program Impact Evaluation in Cooperative Extension: Competencies, Culture and Behaviors

Consider the following definition of ***impact evaluation*** as you answer the questions:

> "... the reportable, quantifiable difference or potential difference a program makes in the lives of real people. It shows a sustainable societal, environmental, and/or economic change," (eXtension Impact Online course, 2017).

Q2.2. What **role/roles do you fill** in the support and/or performance of **impact evaluation** of educational programs within Extension? (Check **all** that apply.)

- ☐ Impact evaluation design.
- ☐ Data collection.
- ☐ Data analysis.
- ☐ Reporting impact evaluation results.
- ☐ Sharing impact evaluation results.
- ☐ Mentoring/assisting those doing impact evaluation
- ☐ None of these roles.
- ☐ Other --

Q10. **Impact Evaluation Competencies --**
This list of competencies/skills is based on lists prepared by professional evaluators (the American Evaluation Association) and skills highlighted in academic literature in the areas of evaluation, impact evaluation, and Extension evaluation.

Please rate each skill on two scales -- first, your level of skill or proficiency (left column) and, second, the importance of the skill/competency to your role(s) in Extension impact evaluation (right column).

	Rate your skill level for each item.				Rate the importance of each item to your Role(s) in Extension impact evaluation.			
	No Skill	Low Skill	Medium Skill	High Skill Level	Not at All Important	Somewhat Unimportant	Somewhat Important	Very Important
1 - Act ethically when planning, doing, and reporting impact evaluation.	○	○	○	○	○	○	○	○
2 - Understand the knowledge base of impact evaluation (theories, models, methods & tools).	○	○	○	○	○	○	○	○
3 - Use evidence of change to make logical evaluation judgments.	○	○	○	○	○	○	○	○
4 - Engage in professional development to learn about impact evaluation.	○	○	○	○	○	○	○	○
5 - Advocate for & support impact evaluation.	○	○	○	○	○	○	○	○
6 - Consider public value of the educational program being evaluated	○	○	○	○	○	○	○	○
7 - Identify types of evidence needed to show program impact (medium and long term changes in social, economic, civic, and environmental conditions).	○	○	○	○	○	○	○	○

Q1.5. **Impact Evaluation Competencies, continued --**

Please rate each skill on two scales -- first, your level of skill or proficiency (left column) and, second, the importance of the skill/competency to your role(s) in Extension impact evaluation (right column).

	Rate your skill level for each item.				Rate the importance of each item to your Role(s) in Extension impact evaluation.			
	No Skill	Low Skill	Medium Skill	High Skill Level	Not at All Important	Somewhat Unimportant	Somewhat Important	Very Important
8 - Design impact evaluation studies to document evidence of change.	○	○	○	○	○	○	○	○
9 - Understand and use appropriate methods for impact evaluation including quantitative, qualitative and mixed methods.	○	○	○	○	○	○	○	○
10 - Create evaluation questions & tools to measure evidence of program impact.	○	○	○	○	○	○	○	○
11 - Collect data from accessible sources using appropriate, ethical procedures.	○	○	○	○	○	○	○	○
12 - Analyze quantitative data appropriately.	○	○	○	○	○	○	○	○
13 - Analyze qualitative data appropriately.	○	○	○	○	○	○	○	○
14 - Interpret impact evaluation findings relevant to the situation/context.	○	○	○	○	○	○	○	○

Q1.6. **Impact Evaluation Competencies, continued--**

Please rate each skill on two scales -- first, your level of skill or proficiency (left column) and, second, the importance of the skill/competency to your role(s) in Extension impact evaluation (right column).

Rate your skill level for each item.					Rate the importance of each item to your Role(s) in Extension impact evaluation.			
No Skill	Low Skill	Medium Skill	High Skill Level		Not at All Important	Somewhat Unimportant	Somewhat Important	Very Important
○	○	○	○	15 - Use impact evaluation evidence/ results to determine program effectiveness.	○	○	○	○
○	○	○	○	16 - Consider context (political, social, economic, environmental) when planning use of impact evaluation results.	○	○	○	○
○	○	○	○	17 - Plan and manage a feasible impact evaluation plan, budget, resources, and timeline.	○	○	○	○
○	○	○	○	18 - Use appropriate technology to manage, analyze, report, and share impact evaluation data & information.	○	○	○	○
○	○	○	○	19 - Encourage constructive interaction and teamwork throughout the impact evaluation.	○	○	○	○
○	○	○	○	20 - Engage and listen to diverse perspectives (users/stakeholders) during impact evaluation.	○	○	○	○
○	○	○	○	21 - Communicate impact evaluation information and results in timely, effective ways.	○	○	○	○

←

→

Q3.1.
Behavior -- Impact Evaluation Behavior & Practice

Items in this section refer specifically to your actual practice of conducting and/or supporting **impact evaluation** as part of your Extension work.

Q3.2.
Please consider **your role(s)** in the ***support of and completion of impact evaluation*** when answering each question below.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I intend to conduct impact evaluation as part of my job.	○	○	○	○	○
Conducting impact evaluation as part of my job is worthwhile.	○	○	○	○	○
Completing effective impact evaluations requires careful planning.	○	○	○	○	○
Conducting impact evaluations as part of my job is interesting.	○	○	○	○	○
People important to me at work think that I should conduct impact evaluations.	○	○	○	○	○
A professional expectation of me is that I conduct impact evaluations.	○	○	○	○	○
It is common to discuss program impact evaluation in my workplace.	○	○	○	○	○

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I am confident that I can conduct impact evaluation.	○	○	○	○	○
I expect to conduct impact evaluation as part of my job.	○	○	○	○	○
I want to conduct impact evaluations as part of my job.	○	○	○	○	○
I plan impact evaluation as part of my program planning process.	○	○	○	○	○
My experiences with impact evaluation in Extension have been positive.	○	○	○	○	○
My experiences with evaluation, in general, have been positive.	○	○	○	○	○
I prefer to create surveys and data collection tools myself rather than using those created by others.	○	○	○	○	○

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I often use surveys or tools created by others (eg. 4-H Common Measures, grant-specific tools).	○	○	○	○	○
Impact evaluation can help us provide better programs and services.	○	○	○	○	○
I have integrated impact evaluation into my program planning process.	○	○	○	○	○
Collaboration/teamwork is encouraged when doing impact evaluation.	○	○	○	○	○
I am required to engage in (or support) impact evaluation.	○	○	○	○	○

←

→

Q3.3. Does your Extension organization have a separate evaluation department or unit (program & evaluation planning)?

- ○ Yes
- ○ No

Q3.7. What level of experience with each of the following evaluation designs have you had when doing and/or supporting program impact evaluation?

	No Experience	Some Experience	Moderate Amount of Experience	Much experience
Pre- and Post- Tests	○	○	○	○
Retrospective Post- then Pre-test (rank what you know after program and what you remember knowing beforehand)	○	○	○	○
Ripple Effect Mapping Evaluation	○	○	○	○
Longitudinal Impact Evaluation Study	○	○	○	○
Case Study	○	○	○	○
Follow-up Impact Study	○	○	○	○
Economic Impact Study	○	○	○	○
Other	○	○	○	○

Q3.8. What portion of the work you have done to support and/or complete impact evaluation is done in the following ways (as a team OR on your own)? (Total must be 100%)

Working on impact evaluation as a team.	100
Working on impact evaluation on my own.	0
Total	100

Q3.9. Which **model**(s) for *educational program and impact evaluation planning* is/are used most often in your Extension place of work? (Choose one or more.)

- ☐ Logic models
- ☐ Bennett's hierarchy
- ☐ Kirkpatrick's levels of change
- ☐ Other

Q3.10. If "other" please indicate the model for educational program & impact evaluation planning used most often in your Extension context. Elaborate if you wish.

Q3.11. What do you need, if anything, to help you do a more effective job of completing **program impact evaluations**? Please describe/elaborate.

Q3.12. What do you see as the biggest challenges or obstacles to improving impact evaluation of programs?

←

→

Q4.1.

Culture (Organizational Culture Regarding Impact Evaluation)

The following questions address the organizational culture within your Extension work context.
Please consider the evaluation culture with respect to **impact evaluation**, specifically.

Q4.2.

For each item, choose the answer that best represents **your opinion** and **your experiences**; not on how you think others may answer, or on Extension's official policy or intent.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
My Extension colleagues and I respect each others' perspectives and opinions.	○	○	○	○	○
We ask each other for information about work issues and activities.	○	○	○	○	○
My Extension colleagues and I continuously look for ways to improve processes, products and services.	○	○	○	○	○
We are given opportunities to think about and reflect on our work.	○	○	○	○	○
We often stop to talk about the pressing work issues we're facing.	○	○	○	○	○
When trying to solve problems, my Extension colleagues and I use a process of working through the problem before identifying solutions.	○	○	○	○	○
There is little competition among me and my colleagues for recognition or rewards.	○	○	○	○	○

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
My Extension colleagues and I operate from a spirit of cooperation, rather than competition.	○	○	○	○	○
Extension colleagues tend to work collaboratively with each other.	○	○	○	○	○
Extension colleagues are more concerned about how our work contributes to the success of the organization than we are about our individual success.	○	○	○	○	○
We face conflict over work issues in productive ways.	○	○	○	○	○
Extension colleagues generally view problems or issues as opportunities to learn.	○	○	○	○	○
Mistakes made by Extension colleagues are viewed as opportunities for learning.	○	○	○	○	○
We continuously ask ourselves how we're doing, what we can do better, and what is working.	○	○	○	○	○
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
My Extension colleagues and I are willing to take risks in the course of our work.	○	○	○	○	○
We are committed to being innovative and forward looking.	○	○	○	○	○
We are confident that mistakes or failures will not affect us negatively.	○	○	○	○	○
My Extension colleagues and I generally trust our supervisors.	○	○	○	○	○
Extension administrators view individuals' capacity to learn as the organization's greatest resource.	○	○	○	○	○
We use data/information to inform our decision-making.	○	○	○	○	○
Asking questions and raising issues about work is encouraged.	○	○	○	○	○
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
My Extension colleagues and I are not afraid to share our opinions even if those opinions are different from the majority.	○	○	○	○	○
I feel safe explaining to others why I think or feel the way I do about an issue.	○	○	○	○	○
We are encouraged to take the lead to initiate change or to try to do something different.	○	○	○	○	○

Extension administrators make decisions after considering the input of those affected.	○	○	○	○	○
In meetings we are encouraged to discuss the values and beliefs that underlie our opinions.	○	○	○	○	○
We are encouraged to offer dissenting opinions and alternative viewpoints.	○	○	○	○	○
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree

←

→

Q5.1.
Demographics
We would like to collect some background information to help us better understand your responses and your work context. This section will complete the survey.

Q5.2. How many **years of experience** do you have working in Extension?

Q5.3. What is your primary program area?

- ○ Agriculture/ Natural Resources (ANR)
- ○ 4-H Youth Development (4-H)
- ○ Family & Consumer Science (FCS)
- ○ Community Resource Development (CRD)
- ○ Other

Q5.5. What type of position do you currently hold within Extension?

- ○ County Agent/Extension Educator/Program Coordinator
- ○ Program Assistant (SnapEd, 4-H, etc.)
- ○ Specialist (state or regional)
- ○ Administrator (district, department, state)
- ○ Other

Q5.6. Other _____

Q5.7.
What is your highest level of academic education?

- ○ HIgh school
- ○ Associate/2-yr degree
- ○ Bachelors
- ○ Masters
- ○ Doctorate
- ○ Other

Q5.8. Have you completed training or staff development in **impact** evaluation, specifically?

- ○ Yes
- ○ Maybe
- ○ No

← →

Q5.9. What types of training and staff development in how to perform ***impact evaluation*** have you had? (Choose all that apply.)

- ☐ On-the-Job experience (learning by doing)
- ☐ Staff development, face-to-face
- ☐ Staff development, online
- ☐ Mentoring from a colleague or administrator
- ☐ Part of an academic course
- ☐ Grant-required; by grant providers
- ☐ Self-directed study
- ☐ Other

Q5.10. Other _____

Q5.11. What is your location (state)?

- Illinois (IL)
- Indiana (IN -- Purdue)
- Iowa (IA)
- Kansas (KS)
- Michigan (MI)
- Minnesota (MN)
- Missouri (MO)
- Nebraska (NE)
- North Dakota (ND)
- Ohio (OH)
- South Dakota (SD)
- Wisconsin (WI)
- Wyoming (WY)

←

→

NDSU NORTH DAKOTA STATE UNIVERSITY

You successfully submitted your survey.
Thank you for your time!

APPENDIX B. LETTER OF CONSENT

This letter of consent was the first page (screen) of the online survey.

NDSU NORTH DAKOTA STATE UNIVERSITY

Q1.1.

Study Introduction & Consent Form

Title of Research Study:
Program Impact Evaluation in Cooperative Extension: Competencies, Culture and Behaviors

This study is being conducted by: Jan Flack, Doctoral Student, and Dr. Claudette Peterson, Adviser.

Why I am being asked to take part in this research study? You are invited to participate in this research because you are a professional Extension Educator/Agent, Specialist, or Administrator within the North Central Region of the United States **who supports and/or engages in doing impact evaluation of educational programs**.

What is the reason for doing the study?
The study explores the factors that influence Extension professionals' efforts to do impact evaluation of educational programs. Results may indicate individual and organizational strengths and challenges for how to support impact evaluation.

What I am being asked to do?
You are invited to complete this online survey. Questions explore your perceptions of your capacity to do impact evaluation and perceptions of the support & assistance for impact evaluation available in your work context. Answers will be kept confidential.

How long will the survey take?
Survey completion will take 20 to 25 minutes. Your answers will be saved and, if you cannot complete the survey the first time, you may access it again using the same link provided initially to complete it.

What are the benefits of participating? You will have an opportunity to reflect on your own impact evaluation experiences and your responses may inform future Extension impact evaluation efforts.

What are the risks and discomforts, if any? Completion of the survey will involve no undue risk or discomfort. You may choose not to answer any question, for any reason. Your answers to all questions will be kept completely confidential and no personal identifying questions will be asked.

Do you have to take part in the study? It is your choice whether or not to participate.

What if I have questions?
This study has been reviewed by the NDSU Institutional Review Board (IRB). If you have any questions about your rights as a human research participant, or if you wish to report a concern, please contact the NDSU Human Research Protection Program at:

- Telephone: 701.231.8995 · Toll Free: 855.800.6717 · Email: ndsu.irb@ndsu.edu
- Mail: NDSU HRPP, 1735 NDSU Research Park Dr., NDSU Dept. 4000, PO Box 6050, Fargo, ND 58108-6050

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/research/irb.

Documentation of Informed Consent:
You are freely making a decision whether to be in this research study. Continuing with this survey means that:

1. You have read and understood this consent form
2. You have decided to be in the study.

>> By clicking on the **Arrow** below you agree to participate in this study.

~ ~ ~ ~ ~ ~ ~ ~ ~ ~

Please note: This survey will be best displayed on a laptop or desktop computer.
Some features may be less compatible for use on a mobile device.

→

APPENDIX C. INVITATION TO PARTICIPATE IN RESEARCH

Date: 9/18 or 9/19 or 9/20/2018

Subject: Impact Evaluation Research Study Invitation to Participate

Dear Extension Professional:
(***Please change the salutation if you wish; please send emails to Extension educators/agents, specialists, & administrators – all are part of the target population for the study.***)

You are invited to participate in a research study exploring impact evaluation in Extension. The study involves completing an online survey. This study is being conducted by me—Jan A. Flack, a doctoral student at North Dakota State University in the department of Education, School of Education. I have Extension experience at the county and state levels in Wisconsin, Minnesota and North Dakota.

This research will explore the skills, knowledge, and attitudes needed for Extension professionals to conduct effective impact evaluation. The study addresses organization evaluation culture and impact evaluation behaviors of Extension county, multi-county, regional, and state Extension educators, specialists, and administrators. Completing the survey will give you a chance to reflect on what influences the effective impact evaluation of educational programs in your context.

Your participation in this research is completely voluntary and does not impact your position in any way. All responses will be kept strictly confidential. You will be asked to provide consent to participate on the first screen of the online survey. The study will take approximately 20 to 25 minutes from start to finish.

Your assistance with this important research will be greatly appreciated. It may provide valuable input for your state and others in the North Central Region. The survey is available at:

https://ndstate.co1.qualtrics.com/jfe/form/SV_cY2yPRbd1BfZUQl

Thank you, in advance, for your time and efforts spent to participate in this important study. If you have any questions regarding this research please contact me at email: jan.flack@ndsu.edu or via phone at 701-388-6779. You may also contact my advisor, Dr. Claudette Peterson, via email at Claudette.peterson@ndsu.edu , if you wish.

Sincerely,
Jan A. Flack
Doctoral Candidate ~ School of Education
North Dakota State University
Fargo, ND 58102

NDSU IRB Approval #HE19042

cc: Dr. Claudette Paterson (Doctoral Adviser)
Dr. [redacted] (MSU State Contact)

APPENDIX D. IRB APPROVAL LETTER

NDSU NORTH DAKOTA STATE UNIVERSITY

September 11, 2018

Dr. Claudette Peterson
School of Education

Re: IRB Determination of Exempt Human Subjects Research:
Protocol #HE19042, "Program Impact Evaluation in Cooperative Extension: Behaviors, Competencies, and Culture"

Co-investigator(s) and research team: Jan A. Flack, Chris Ray
Date of Exempt Determination: 9/11/2018 Expiration Date: 9/10/2021
Study site(s): varied, online
Sponsor: n/a

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 revised protocol submission (received 9/11/2018).

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.

APPENDIX E. CODE BOOK ~ STUDY VARIABLES

This appendix is found on the following pages. It lists the item and variable names used in data analysis, the actual items and original items if the item was adapted from one previously used, the source of the item, and the construct being measured.

Code Book – Study Variables

Impact Evaluation Competency Items

Variable name	Item used in study	Original item	Source(s)	Construct
comp_s	Composite of competency by perceived skill items.	See Appendix G for Stata code for this variable.	Item sources.	Competency by Skill
comp_i	Composite of competency by perceived importance items.	See Appendix G for Stata code for this variable.	Item source.	Competency by Importance
c_pro_s1 (q2_4_1_1)	Act ethically when planning, doing, and reporting impact evaluation.	2.1 Acts ethically in the conduct of evaluation inquiry.	AEA Evaluator Competencies, 2018	Competency Professional Practice
c_pro_s2 (q2_4_1_2)	Understand the knowledge base of impact evaluation (theories, models, methods & tools).	2.1 Understands the knowledge base of evaluation (theories, models, types, methods and tools) --- 1.1 Knows and applies program evaluation foundations that ground and guide professional practice (e.g., standards, guidelines, principles, competencies, approaches, and theories).	Canadian Eval Competencies --- AEA Evaluator Competencies	Competency – Professional Practice
c_mth_s1 (q2_4_1_3)	Use evidence of change to make logical evaluation judgments.	2.14 Uses evidence and interpretations to draw conclusions, making judgments and recommendations when appropriate.	AEA Evaluator Competencies	Competency Methods
c_pro_s3 (q2_4_1_4)	Engage in professional development to learn about impact evaluation.	1.7 Pursues ongoing professional development to deepen reflective practice, stay current, and build connections. 1.4 Engages in ongoing professional development to extend personal learning and growth.	AEA Evaluator Competencies	Competency Professional Practice
c_pro_s4 (q2_4_1_5)	Advocate for & support impact evaluation.	1.9 Advocates for the field of evaluation and its value.	AEA Evaluator Competencies	Competency ProfPractice

c_ctx_s3 (q2_4_1_6)	Consider public value of the educational program being evaluated.	Original item – adapted from Canadian item re: context-specific measurement challenges -- 3.8 Applies evaluation competencies to organization and program measurement challenges.	Canadian Evaluator Competencies	Competency – Context

Variable name	Item used in study	Original item	Source(s)	Construct
c_pro_i1 (q2_4_2_1)	Act ethically when planning, doing, and reporting impact evaluation.	See q2_4_1_1	AEA Evaluator Competencies	Competency – ProfPractice
c_pro_i2 (q2_4_2_2)	Understand the knowledge base of impact evaluation (theories, models, methods & tools).	See q2_4_1_2	AEA Evaluator Competencies	Competency – ProfPractice
c_mth_i1 (q2_4_2_3)	Use evidence of change to make logical evaluation judgments.	See q2_4_1_3	AEA Evaluator Competencies	Competency Methods
c_pro_i3 (q2_4_2_4)	Engage in professional development to learn about impact evaluation.	See q2_4_1_4	AEA Evaluator Competencies	Competency – ProfPractice
c_pro_i4 (q2_4_2_5)	Advocate for & support impact evaluation.	See q2_4_1_5	AEA Evaluator Competencies	Competency – ProfPractice
c_ctx_i3 (q2_4_2_6)	Consider public value of the educational program being evaluated.	See q2_4_1_6	AEA Evaluator Competencies	Competency – Context
c_mth_s2 (q2_4_1_7)	Identify types of evidence needed to show program impact (medium- and long-term changes in social, economic, civic, and environmental conditions).	Original item based on – 2.9 Develops reliable and valid measures/tools	AEA Evaluator Competencies	Competency – Methods
c_mth_s3 (q2_5_1_1)	Design impact evaluation studies to document evidence of change.	Original item based on – 2.9 Develops reliable and valid measures/tools	AEA Evaluator Competencies	Competency – Methods
c_mth_s4 (q2_5_1_2)	Understand and use appropriate methods for impact evaluation including quantitative, qualitative and mixed methods.	Adapted from – 2.7 Defines evaluation methods (quantitative, qualitative or mixed)	AEA Evaluator Competencies	Competency – Methods

c_mth_s5 (q2_5_1_3)	Create evaluation questions & tools to measure evidence of program impact.	Original item based on – 2.6 Designs credible and feasible studies that address evaluation purposes and questions. and 2.9 Develops reliable and valid measures/tools	AEA Evaluator Competencies	Competency – Methods
c_mth_s6 (q2_5_1_4)	Collect data from accessible sources using appropriate, ethical procedures.	2.9 Collects data using sound and credible procedures. (2017 draft)	AEA Evaluator Competencies	Competency – Methods

Variable name	Item used in study	Original item	Source(s)	Construct
c c_mth_s7 (q2_5_1_5)	Analyze quantitative data appropriately.	Analyze quantitative data	Preskill & Boyle, 2008	Competencies– Methods
c_mth_s8 (q2_5_1_6)	Analyze qualitative data appropriately.	Analyze qualitative data	Preskill & Boyle, 2008	Competencies – Methods
c_mth_i2 (q2_4_2_7)	Identify types of evidence needed to show program impact (medium- and long-term changes in social, economic, civic, and environmental conditions).	Same as at left.	AEA Evaluator Competencies	Competencies –Methods
c_mth_i3 (q2_5_2_1)	Design impact evaluation studies to document evidence of change.	Same as at left.	AEA Evaluator Competencies	Competencies – Methods
c_mth_i4 (q2_5_2_2)	Understand and use appropriate methods for impact evaluation including quantitative, qualitative and mixed methods.	Same as at left.	AEA Evaluator Competencies	Competencies – Methods
c_mth_i5 (q2_5_2_3)	Create evaluation questions & tools to measure evidence of program impact.	Same as at left.	AEA Evaluator Competencies	Competency – Methods
c_mth_i6 (q2_5_2_4)	Collect data from accessible sources using appropriate, ethical procedures.	Same as at left.	AEA Evaluator Competencies	Competency – Methods
c_mth_i7 (q2_5_2_5)	Analyze quantitative data appropriately.	Same as at left.	AEA Evaluator Competencies	Competency – Methods

c_mth_i8 (q2_5_2_6)	Analyze qualitative data appropriately.	Same as at left.	AEA Evaluator Competencies	Competency – Methods
c_ctx_s1 (q2_5_1_7)	Interpret impact evaluation findings relevant to the situation/context.	2.13 Interprets findings/results in context.	AEA Evaluator Competencies	Competency – Context
c_mth_s9 (q2_6_1_1)	Use impact evaluation evidence/ results to determine program effectiveness.	Original item adapted from – 2.14 Analyzes and interprets data 2.15 Draws conclusions & makes recommendations	AEA Evaluator Competencies	Competency – Methods

Variable name	Item used in study	Original item	Source(s)	Construct
c_ctx_s2 (q2_6_1_2)	Consider context (political, social, economic, environmental) when planning use of impact evaluation results.	3.2 Examines organizational, political, community and social contexts 3.6 Attends to issues of evaluation use in context, including the information needs of intended users.	AEA Evaluator Competencies	Competency – Context
c_ctx_i1 (q2_5_2_7)	Interpret impact evaluation findings relevant to the situation/context.	2.13 Interprets findings/results in context. 2.11 Interprets findings/results and draws conclusions by identifying possible meanings in context.	AEA Evaluator Competencies	Competency – Context
c_mth_i9 (q2_6_2_1)	Use impact evaluation evidence/ results to determine program effectiveness.	See q2_6_1_1	AEA Evaluator Competencies	Competency – Methods
c_ctx_i2 (q2_6_2_2)	Consider context (political, social, economic, environmental) when planning use of impact evaluation results.	See q2_6_1_2	AEA Evaluator Competencies	Competency – Context
c_pam_s1 (q2_6_1_3)	Create and manage a feasible impact evaluation plan, budget, resources, and timeline.	4.1 Negotiates and manages a feasible evaluation plan, budget, resources, and timeline.	AEA Evaluator Competencies	Competency – PlngMngmnt
c_pam_s2 (q2_6_1_4)	Use appropriate technology to manage, analyze, report, and share impact evaluation data & information.	4.10 Uses technology appropriately to support and manage the evaluation.	AEA Evaluator Competencies	Competency—PlngMngmt

c_int_s1 (q2_6_1_5)	Encourage constructive interaction and teamwork throughout the impact evaluation.	5.6 Facilitates constructive interaction among those involved in the evaluation, . 5.7 Applies teamwork skills for collaborative endeavors in evaluation.	AEA Evaluator Competencies	Competency – Interpersonal
c_int_s2 (q2_6_1_6)	Engage and listen to diverse perspectives (users/stakeholders) during impact evaluation.	5.3 Listens to understand and engages diverse perspectives in evaluation.	AEA Evaluator Competencies	Competency – Interpersonal
c_int_s3 (q2_6_1_7)	Communicate impact evaluation information and results in timely, effective ways.	5.6 Communicates in meaningful ways throughout the evaluation (written, verbal, visual, etc.).	AEA Evaluator Competencies	Competency – Interpersonal
c_pam_i1 (q2_6_2_3)	Create and manage a feasible impact evaluation plan, budget, resources, and timeline.	See q2_6_1_3	AEA Evaluator Competencies	Competency – PlngMgmt
c_pam_i2 (q2_6_2_4)	Use appropriate technology to manage, analyze, report, and share impact evaluation data & information.	See q2_6_1_4	AEA Evaluator Competencies	Competency – PlngMgmt
c_int_i1 (q2_6_2_5)	Encourage constructive interaction and teamwork throughout the impact evaluation.	See q2_6_1_5	AEA Evaluator Competencies	Competency – Interpersonal
c_int_i2 (q2_6_2_6)	Engage and listen to diverse perspectives (users/stakeholders) during impact evaluation.	See q2_6_1_6	AEA Evaluator Competencies	Competency – Interpersonal
c_int_i2 (q2_6_2_7)	Communicate impact evaluation information and results in timely, effective ways.	See q2_6_1_7	AEA Evaluator Competencies	Competency – Interpersonal

Theory of Planned Behavior (TPB) Items

Variable name	Item used in study	Original item	Source(s)	Construct
att	Composite tpb attitude variable.	See individual items.	See items.	Attitude.
norm	Composite tpb subjective norm variable.	See individual items.	See items.	Subjective norm.
pbc	Composite tpb perceived behavioral control variable.	See individual items.	See items.	Perceived behavioral control

bint	Composite tpb behavioral intent variable.	See individual items.	See items coded bint below.	Behavioral intent
tpb_bint1 (q3_2_1)	I intend to conduct impact evaluation as part of my job.	I intend to conduct impact evaluation as part of my job. (Behavioral Intention, Vengrin, Q28-1.)	Vengrin, 2016	Behavioral Intent
tpb_att1 (q3_2_2)	Conducting impact evaluation as part of my job is worthwhile.	Conducting evaluation as part of my job is worthless … useful. (Attitude, Vengrin, Q7-1.)	Vengrin, 2016	Attitude
tpb_att2 (q3_2_3)	Completing effective impact evaluations requires careful planning.	Similar item from Smith – My using EBCT teaching strategies in my classroom/ clinical rotation this school year for critical thinking development in nursing students would demand more planning time.	***Original Item*** Smith (2015); Wells-Marshall (2012); Montaño & Kasprzyk (2015); Ajzen (2006)	Attitude

Variable name	**Item used in study**	**Original item**	**Source(s)**	**Construct**
tpb_att3 (q3_2_4)	Conducting impact evaluations as part of my job is interesting.	Conducting evaluations as part of my job is interesting. (Attitude, Vengrin Q8-1)	Vengrin, 2016	Attitude
tpb_norm1 (q3_2_5)	People important to me at work think that I should conduct impact evaluations.	People who are important to me want me to conduct evaluations. (Subj Norm, Vengrin Q15-1)	Vengrin, 2016; Smith, 2015; Teo & Lee, 2010	SubjNorm
tpb_norm2 (q3_2_7)	It is common to discuss program impact evaluation in my workplace.	It is common to discuss program evaluations in my workplace. (Subj Norm, Vengrin Q13-1.)	Vengrin, 2016	Subj Norm
tpb_norm3 (q3_2_6)	A professional expectation of me is that I conduct impact evaluations.	It is expected of me that I conduct evaluations.	Vengrin, 2016	Subjective Norm
tpb_pbc1 (q3_2_8)	I am confident that I can conduct impact evaluation.	I am confident I can conduct evaluations if I want to. (PercBehCtrl, Vengrin Q34-1.)	Vengrin, 2016	PercBehCtrl

tpb_norm4 (q3_2_9)	I expect to conduct impact evaluation as part of my job.	It is expected of me that I conduct evaluations as part of my job. (Subj Norm, Vengrin, Q12-1)	Vengrin, 2016	SubjNorm
tpb_bint2 (q3_2_10)	I want to conduct impact evaluations as part of my job.	I want to conduct evaluation as part of my job. (Behavioral Intention, Vengrin, Q3-4.)	Vengrin, 2016	Beh Intent
tpb_bint3 (q3_2_11)	I plan impact evaluation as part of my program planning process.	Original item based on literature (see sources).	Wells-Marshall, 2012; Montaño & Kasprzyk, 2015; Ajzen, 2006	Beh Intent
tpb_att4 (q3_2_12)	My experiences with impact evaluation in Extension have been positive.	Original item based on literature (see sources).	Taut & Brauns, 2003; Ajzen, 2006	Attitude
tpb_att5 q3_2_13	My experiences with evaluation, in general, have been positive.	Original item based on literature (see sources).	Taut & Brauns, 2003; Ajzen, 2006	Attitude
tpb_pbc2 q3_2_14	I prefer to create surveys and data collection tools myself rather than using those created by others.	Original item based on literature (see sources).	Montaño & Kasprzyk, 2015; Ajzen, 2006	PercBeh Ctrl

Variable name	**Item used in study**	**Original item**	**Source(s)**	**Construct**
tpb_pbc4 q3_2_15	I often use surveys or tools created by others (eg. 4-H Common Measures, grant-specific tools).	Original item based on literature (see sources).	Smith, 2015; Ajzen, 2006	PercBeh Ctrl
tpb_att6 q3_2_16	Impact evaluation can help us provide better programs and services.	Evaluation helps (or would help) us provide better programs, processes, products and services. (Preskill & Torres, Q. 74.)	Preskill & Torres, 2000	Attitude
tpb_bint4 q3_2_17	I have integrated impact evaluation into my program planning process.	Original item based on literature (see sources).	Wells-Marshall, 2012; Ajzen, 2006	Behavioral Intention
tpb_norm5 q3_2_18,	Collaboration/teamwork is encouraged when doing impact evaluation.	Original item based on literature (see sources).	American Evaluation Assn. (2018); Wells-Marshall	SubjNrm

			(2012); Ajzen (2006)	
tpb_pbc3 q3_2_19	I am required to engage in (or support) impact evaluation.	Q38-1 – Whether or not I conduct evaluations is entirely up to me. (PercBehContrl) Vengrin, 2016	Vengrin, 2016	PercBeh Ctrl

Theory of Planned Behavior (TPB) – Actual Impact Evaluation (IE) Behaviors

Variable Name	Item used in study	Original item	Source(s)	Construct
actbx (q3_7)	Composite variable for actual impact evaluation behaviors. What level of experience with each of the following evaluation designs have you had when doing and/or supporting program impact evaluation? No Experience (1) / Some Experience (2) / Moderate Amount of Experience (4) / Much experience (5)	Original item based on categories of actual impact evaluation behaviors found in literature cited for individual sub-items shown in this table. .	See sources provided for each of the seven categories of actual impact evaluation behaviors reported.	Actual IE Behaviors
tpb_actbx1 (q3_7_1)	Pre- and Post- Tests (1)	Pre- and Post- Tests	Roche (1999); Leeuw, & Vaessen, (2009)	Actual IE Behaviors
tpb_actbx2 (q3_7_2)	Retrospective Post- then Pre-test (indicate what you know after a program and what you remember knowing beforehand) (2)	Retrospective Post- then Pre-test	Huddleston-Casas, Danes, & Boyce (1999).	Actual IE Behaviors
tpb_actbx3 (q3_7_3)	Ripple Effect Mapping Evaluation (3)	Ripple effect mapping.	Kollock, Flage, Chazdon, Paine, & Higgins (2012).	Actual IE Behaviors
tpb_actbx4 (q3_7_4)	Longitudinal Impact Evaluation Study (4)	Longitudinal Impact Evaluation Study	Kreber, Brook, & Policy, (2001); Leeuw, & Vaessen, (2009)	Actual IE Behaviors

tpb_actbx5 (q3_7_5)	Case Study (5)	Case study	Zaleski, Kushner, Pratsch, Clugston, & Jones, (2014).	Actual IE Behaviors
tpb_actb6 (q3_7_6)	Follow-up Impact Study (6)	Follow-up impact study	Arts, & Morrison-Saunders, (Eds.). (2012).	Actual IE Behaviors
tpb_actbx7 (q3_7_7)	Economic Impact Study	Economic impact study	Howes, 1992; Leeuw & Vaesen, 2009.	Actual IE Behaviors

Org/Impact Evaluation (IE) Culture Items – All from Preskill & Torres, 1999 (Also used by Vengrin, 2016)

Variable name	Item used in study	Original item	Source(s)	Construct
cult	Composite variable for all IE culture items.	Calculated from means of all IE culture items.	Preskill & Torres, 1999	Culture (IE)
cult_cps1 (q4_2_1)	My Extension colleagues and I respect each other's perspectives and opinions.	Employees respect each other's perspectives and opinions.	Preskill & Torres, 1999. Reliability info: Preskill, Torres, & Martinez-Papponi, 1999.	Culture – *Collaboration and Problem Solving*
cult_cps2 (q4_2_2)	We ask each other for information about work issues and activities.	Employees ask each other for information about work issues and activities.	Preskill & Torres, 1999	Culture – *Collaboration and Problem Solving*

Variable name	Item used in study	Original item	Source(s)	Construct
cult_cps3 (q4_2_3)	My Extension colleagues and I continuously look for ways to improve processes, products and services.	Employees continuously look for ways to improve processes, products and services.	Preskill & Torres, 1999	Culture – *Collaboration and Problem Solving*
cult_cps4 (q4_2_4)	We are given opportunities to think about and reflect on our work.	Employees are provided opportunities to think about and reflect on our work.	Preskill & Torres, 1999	Culture – *Collaboration and Problem Solving*

cult_cps5 (q4_2_5)	We often stop to talk about the pressing work issues we're facing.	Employees often stop to talk about the pressing work issues we're facing.	Preskill & Torres, 1999	Culture – *Collaboration/ ProblemSolving*
cult_cps6 (q4_2_6)	When trying to solve problems, my Extension colleagues and I use a process of working through the problem before identifying solutions.	When trying to solve problems, employees use a process of working through the problem before identifying solutions.	Preskill & Torres, 1999	Culture – *Collaboration/ ProblemSolving*
cult_cps7 (q4_2_7)	There is little competition among me and my colleagues for recognition or rewards.	There is little competition among employees for recognition or rewards.	Preskill & Torres, 1999	Culture – *Collaboration/ ProblemSolving*
cult_cps8 (q4_2_8)	My Extension colleagues and I operate from a spirit of cooperation, rather than competition.	Employees operate from a spirit of cooperation, rather than competition.	Preskill & Torres, 1999	Culture – *Collaboration/ ProblemSolving*
cult_cps9 (q4_2_9)	Extension colleagues tend to work collaboratively with each other.	Employees tend to work collaboratively with each other.	Preskill & Torres, 1999	Culture – *Collaboration/ ProblemSolving*
cult_cps10 (q4_2_10)	Extension colleagues are more concerned about how our work contributes to the success of the organization than we are about our individual success.	Employees are more concerned about how their work contributes to the success of the organization than they are about our individual success.	Preskill & Torres, 1999	Culture – *Collaboration/ ProblemSolving*
cult_cps11 (q4_2_11)	We face conflict over work issues in productive ways.	Employees face conflict over work issues in productive ways.	Preskill & Torres, 1999	Culture – *Collaboration/ ProblemSolving*
cult_cps12 (q4_2_12)	Extension colleagues generally view problems or issues as opportunities to learn.	Employees generally view problems or issues as opportunities to learn.	Preskill & Torres, 1999	Culture – *Collaboration/ ProblemSolving*

Variable name	**Item used in study**	**Original item**	**Source(s)**	**Construct**
cult_rt1 q4_2_13	Mistakes made by Extension colleagues are viewed as opportunities for learning.	Mistakes made by employees are viewed as opportunities for learning.	Preskill & Torres, 1999	Culture- Risk Taking
cult_rt2 (q4_2_14)	We continuously ask ourselves how we're doing, what we can do better, and what is working.	Employees continuously ask themselves how they're doing, what they can do better, and what is working.	Preskill & Torres, 1999	Culture- Risk Taking

cult_rt3 (q4_2_15)	My Extension colleagues and I are willing to take risks in the course of our work.	Employees are willing to take risks in the course of their work.	Preskill & Torres, 1999	Culture- Risk Taking
cult_rt4 (q4_2_16)	We are committed to being innovative and forward looking.	Employees are committed to being innovative and forward looking.	Preskill & Torres, 1999	Culture- Risk Taking
cult_rt5 (q4_2_17)	We are confident that mistakes or failures will not affect us negatively.	Employees are confident that mistakes or failures will not affect them negatively.	Preskill & Torres, 1999	Culture- Risk Taking
cult_pdm1 (q4_2_18)	My Extension colleagues and I generally trust our supervisors.	Employees generally trust their managers or supervisors.	Preskill & Torres, 1999	Culture- *Participatory Decision Making*
cult_pdm2 (4_2_19)	Extension administrators view individuals' capacity to learn as the organization's greatest resource.	Managers and supervisors view individuals' capacity to learn as the organization's greatest resource.	Preskill & Torres, 1999 Preskill & Torres, 1999	Culture- *Participatory Decision Making*
cult_pdm3 (q4_2_20)	We use data/information to inform our decision-making.	Employees use data/information to inform their decision-making.	Preskill & Torres, 1999	Culture- *Participatory Decision Making*
cult_pdm84 (q4_2_21)	Asking questions and raising issues about work is encouraged.	Asking questions and raising issues about work is encouraged.	Preskill & Torres, 1999	Culture- *Participatory Decision Making*
cult_pdm5 (q4_2_22)	My Extension colleagues and I are not afraid to share our opinions even if those opinions are different from the majority.	Employees are not afraid to share their opinions even if those opinions are different from the majority.	Preskill & Torres, 1999	Culture- *Participatory Decision Making*
cult_pdm6 (q4_2_23)	I feel safe explaining to others why I think or feel the way I do about an issue.	I feel safe explaining to others why I think or feel the way I do about an issue.	Preskill & Torres, 1999	Culture- *Participatory Decision Making*
cult_pdm7 (q4_2_24)	We are encouraged to take the lead to initiate change or to try to do something different.	Employees are encouraged to take the lead in initiating change or in trying to do something different.	Preskill & Torres, 1999	Culture- *Participatory Decision Making*

cult_pdm8 (q4_2_25)	Extension administrators make decisions after considering the input of those affected.	Managers and supervisors make decisions after considering the input of those affected.	Preskill & Torres, 1999	Culture-*Participatory Decision Making*
cult_pdm9 (q4_2_26)	In meetings we are encouraged to discuss the values and beliefs that underlie our opinions.	In meetings employees are encouraged to discuss the values and beliefs that underlie their opinions.	Preskill & Torres, 1999	Culture-*Participatory Decision Making*
cult_pdm10 (q4_2_27)	We are encouraged to offer dissenting opinions and alternative viewpoints.	Employees are encouraged to offer dissenting opinions and alternative viewpoints.	Preskill & Torres, 1999	Culture-*Participatory Decision Making*

Demographic Variables

Variable name	Item used in study	Original item	Source(s)	Construct
Individual Demographic Factors				
train	Training specifically in IE? Yes/Maybe/No	Same item as used by Vengrin to assess training in general evaluation; used to assess IE training in this study.	Vengrin (2016)	Training
trainex	Indicate types of IE training – see Appendix A, question 5.9 for the list of training types included.	Yes – original item. Types of training taken from others' research.	Wells-Marshal, 2012; Berrio, ; Vengrin, 2016	Types of training
progarea	Which is your major program area? __ANR __4-H __ FCS __ CRD __ other See Appendix A. Q5.3 for details.	Program areas taken from others' research and updated based on expert reviewer feedback.	Expert reviewers; Berrio, 1999; Vengrin, 2016	Program Area
position	What is your position in Extension? See Appendix A, Q5.5 for list.	Taken from others" research	Berrio, 1999; Vengrin, 2016	Type of position
exper	Years of experience in the Extension service: _____ (text box)	Taken from similar items used in others' research.	Berrio, 1999; Vengrin, 2016	Experience (years in Ext.)
edlevel	What is your highest level of academic education? See Appendix A., Q5.7 for list.	Similar to items from other research and updates based on expert reviewer feedback.	Vengrin, 2016; Berrio, 1999	Education level

num_roles	Which role(s) do you fill in impact evaluation? See Appendix A., question 2.1 for list of roles.	Yes. Roles gleaned from research; item concept from role theory.	Van der Horst, 2016	Number of IE roles
Organizational Demographic Factors				
support	Does your organization have a separate evaluation/program planning unit? __yes __no	Original item—Q3.3. Support measured in various ways in other research; this study targeted organizational structure.	Lamm, 2011; Silliman, et al., 2016a & 2016b	Org Support -- structure
location	Indicate the state where you work: NCR state list. See Appendix A. Q5.11. for list of states.	Twelve states in the North Central Region (NCR).	Extension, NCR.	Location (state)
propteam	Proportion of impact eval done with teamwork. Individual ←---25-----50 ----75----→ Teamwork	Original item. Based on literature and researcher experience.	Roche, 1999; Nichols, et al., 2015.	Proportion of Teamwork
num_iemodels	Program planning models used? __Logic model, __ Kirkpatrick, __Bennett's Hierarchy, __Other.	Original item – Q3.9.Similar to items used in other research.	Workman 2012; Vengrin, 2016	Type(s) of IE models used

APPENDIX F. STATA CODE FOR CALCULATING VARIABLES

Variable/Factor	Stata Code Employed to Create Variable
Theory of Planned Behavior (TPB) Attitude → **att**	// Composite score for Theory of Planned Behavior – attitude egen att = rowmean(tpb_att*)
TPB Subjective norm → **norm**	// Composite score for Theory of Planned Behavior - subjective norms egen norm = rowmean(tpb_norm*)
TPB Perceived Behavior Control → **pbc**	// Composite score for Theory of Planned Behavior - perceived behavior control egen pbc = rowmean(tpb_pbc*)
TPB Behavioral Intention →**bint**	// Composite score for Theory of Planned Behavior - behavioral intention egen bint = rowmean(tpb_bint*)
TPB Actual IE Behavior →**actbx**	// Composite score for Theory of Planned Behavior - actual behavior egen actbx = rowmean(tpb_actbx*)
Organization Impact Evaluation (IE) Culture →**cult**	// Culture - overall composite egen cult = rowmean(cult_cps cult_rt cult_pdm)
Competency by Perceived Skill → **comp_s**	// Competency skill - overall composite egen comp_s = rowmean(c_*_s)
Competency by Perceived Importance → **comp_i**	// Competency importance - overall composite egen comp_i = rowmean(c_*_i)
Training (Y/N) → **train** Note: Training "maybe" was recoded as "no" for this variable.	// Impact eval training (yes-->1; maybe,no-->0) gen train = 1 if ietrain==1 replace train = 0 if inlist(ietrain,2,3)
Years of Experience → **exper** Note: Years of experience were recoded from text strings to numerical data.	// Convert string to numerical destring exper, ignore("+") replace
Position recoded → **position**	// Fix response code numbering recode position (5=2) (2=3) (3=4) (4=5)

Education levels recoded → **edlevel**	// Fix response code numbering recode edlevel (5=1) (6=2) (1=3) (2=4) (3=5) (4=6)

Number of Roles – **num_roles**	// Roles
	gen role1 = strmatch(roles, "*1*")
	gen role2 = strmatch(roles, "*2*")
	gen role3 = strmatch(roles, "*3*")
	gen role4 = strmatch(roles, "*4*")
	gen role5 = strmatch(roles, "*5*")
	gen role6 = strmatch(roles, "*7*") // 7 --> 6
	gen role7 = strmatch(roles, "*8*") // 8 (other) --> 7
	gen role8 = strmatch(roles, "*6*") // 6 (none) --> 8
	egen num_roles = rowtotal(role1-role7) // 8 (none) omitted from count

Number of Impact Evaluation Models Used → **num_iemodels**	// Impact evaluation models used gen iemodel1 = strmatch(iemodels, "*1*") gen iemodel2 = strmatch(iemodels, "*2*") gen iemodel3 = strmatch(iemodels, "*3*") gen iemodel4 = strmatch(iemodels, "*4*") egen num_iemodels = rowtotal(iemodel1-iemodel4)

IE Culture Variables Used for Parceling Data**

Variable/Factor	**Stata Code Employed to Create Value for Variable**
Culture – collaboration & problem solving → **cult_cps**	// Composite score for Culture - collaboration and problem solving egen cult_cps = rowmean(cult_cps*)
Culture – risk taking → **cult_rt**	// Composite score for Culture - risk taking egen cult_rt = rowmean(cult_rt*)
Culture – participatory decision making → **cult_pdm**	// Composite score for Culture - participatory decision making egen cult_pdm = rowmean(cult_pdm*)

Note. ** Parceling of data was employed in model estimation; parceled data variables were not found to be significant in the model and were not included in the final structural model.

APPENDIX G. STATA CODE FOR SEM MODEL ESTIMATION

Model Being Estimated	Stata Code for SEM Estimation of a Model

Initial Model	// Initial model with separate competency skill and importance and demographics
	sem ///
	(actbx <- bint comp_s) ///
	(bint <- att norm comp_s) ///
	(att <- cult comp_i comp_s) ///
	(norm <- cult comp_i comp_s) ///
	(pbc <- cult comp_i comp_s) ///
	(cult <- propteam edlevel num_roles num_iemodels train exper) ///
	(comp_s <- propteam edlevel num_roles num_iemodels train exper) ///
	(comp_i <- propteam edlevel num_roles num_iemodels train exper), ///
	cov(_oex e.pbc*e.norm e.norm*e.att e.att*e.pbc
	e.comp_s*e.comp_i) nomeans stand

Final Structural Model	// Final model with separate competency skill and importance and demographics
	sem ///
	(actbx <- bint comp_s) ///
	(bint <- att norm comp_s) ///
	(att <- cult comp_i comp_s) ///
	(norm <- cult comp_i comp_s) ///
	(pbc <- cult comp_i comp_s) ///
	(cult <- propteam) ///
	(comp_s <- propteam edlevel num_roles train) ///
	(comp_i <- propteam edlevel num_roles), ///
	cov(_oex e.pbc*e.norm e.norm*e.att e.att*e.pbc ///
	e.comp_s*e.comp_i e.comp_s*e.cult e.cult*e.comp_i) nomeans stand

Calculate Goodness of Fit Statistics for A Model

Value to Be Calculated	Stata Code
Calculate Goodness of Fit Statistics (gof). Done for each model estimated.	estat gof, stats(all)

Calculate Direct, Indirect and Total Effects for A Model	
Value to Be Calculated	**Stata Code**
Calculate Total Effects (teffects) for model; include standardized values. This command also calculates indirect and direct effects.	estat teffects, stand

Calculate Modification Indices for A Model	
Value to Be Calculated	**Stata Code**
Calculate Modification Indices (mi), chi-squared, and related fit indices. Done for each estimated model.	estat mi, minchi2(10)

Calculate R^2 for A Model	
Value to Be Calculated	**Stata Code**
Calculate R^2 to assess equation-level goodness of model fit. Done for initial and final models.	estat eqgof

Covariance Matrix for Final Structural Model	
Values to Be Calculated	**Stata Code**
Display covariance matrix for final structural model.	corr comp_s comp_i cult att norm pbc bint actbx propteam edlevel num_roles train, cov