

Utah State University

DigitalCommons@USU

Making Connections

Empower Teaching Open Access Book Series

5-15-2023

Chapter 14- The Mentoring Program as a Research Project

David Law

Utah State University, david.law@usu.edu

Nicole Vouvalis

Utah State University, nicole.vouvalis@usu.edu

Andy Harris

Utah State University, andy.harris@usu.edu

Jim LaMuth

Utah State University, jim.lamuth@usu.edu

Follow this and additional works at: <https://digitalcommons.usu.edu/makingconnections>



Part of the [Education Commons](#)

Recommended Citation

Law, David; Vouvalis, Nicole; Harris, Andy; and LaMuth, Jim, "Chapter 14- The Mentoring Program as a Research Project" (2023). *Making Connections*. Paper 16.

<https://digitalcommons.usu.edu/makingconnections/16>

This Chapter is brought to you for free and open access by the Empower Teaching Open Access Book Series at DigitalCommons@USU. It has been accepted for inclusion in Making Connections by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



THE MENTORING PROGRAM AS A RESEARCH PROJECT

David Law; Nicole Vouvalis; Andy Harris; and Jim LaMuth

Abstract

Chapter 14, “The Mentoring Program as a Research Project,” helps stakeholders, program coordinators, and researchers distinguish the differences and similarities between program evaluation and program research. If stakeholders choose to include program research, they will need approval from their university’s institutional review board (IRB). Therefore, the second section of this chapter helps stakeholders navigate the IRB. The third section of this chapter describes how theoretical frameworks, operational definitions of mentoring, and methodological designs factor into mentoring programs that contain research. While all formal mentoring programs in academia should include theoretical frameworks, operational definitions, and sound methodology, many do not. The third section of this chapter highlights the interconnectedness between theory, definitions, methods, and measurements. The fourth and final section provides examples of measurements that can be used. Some of these measurements may be used for both evaluative and research purposes.

Introduction

As described in Chapter 13 by Lunsford, modern mentoring programs are expected to be evaluated. When stakeholders of formal mentoring programs in academia give input into the design/redesign, implementation, and evaluation of their mentoring program, they may also consider whether they want to include program research. For example, stakeholders could ponder such questions as Do we want our program's findings to contribute to generalizable knowledge in the academic mentoring field, or is the intent to provide information for and about our program only? Are we answering questions or testing hypotheses, or are we simply assessing the effectiveness of our program?

Stakeholders' answers to such questions provide insight regarding whether or not they should include program research as they design their program. Chapter 14 guides stakeholders who choose to include a research program in addition to their evaluation program. Program evaluation and program research have different and distinctive processes. However, some program evaluation and research activities, such as data collection, may have similar processes but different purposes. Thus, the differences and similarities between program evaluation and research can confuse novice program coordinators and other university leaders. We offer the following four sections to provide clarity to the novice program coordinator and help them make an informed choice of including program research in addition to program evaluation. The first section helps the reader distinguish between what is considered program evaluation and what is considered program research in academic mentoring programs. If stakeholders choose to include research, they will need approval from their university's institutional review board (IRB). Therefore, the second section of this chapter is to help the reader navigate the IRB. A long-standing shortcoming in formal mentoring programs in academia is a need for methodological rigor. The third section describes ways to improve methodological rigor to help program research contribute to the science of mentoring. The third section also highlights the interconnectedness of theories, operational definitions, research variables, and measurements. The fourth section of this chapter describes possible measurements to consider for research. Some of these measurements can be used for both evaluative and research purposes.

Differences and Similarities Between Program Evaluation and Program Research

An evaluation plan is critical for every formal mentoring program in academia. Figure 7.1 in Chapter 7 visually displays how evaluation informs all aspects of the mentoring program. Most formal mentoring programs have an evaluation plan; however, only some programs include a research component. While stakeholders (including university administrators and program coordinators) may consult with the IRB, ultimately, the IRB will determine whether program activities fall within the research category. For example, the National Institute of Health's (NIH) website on human research design (<https://irbo.nih.gov>) provides guidelines for methodological design, selecting subjects, and publicizing results. Still, it does not necessarily differentiate evaluation from research because some activities can occur in both. The NIH also has provided the following guidance to provide clarity between program evaluation and program research. We start with the following definition of program evaluation: "The Centers for Disease Control defines program evaluation as a systematic method for collecting, analyzing, and using data to examine the effectiveness and efficiency of programs and, as significantly, to contribute to continuous program improvement" (Centers for Disease Control

and Prevention 2022). This website provides critical concepts to help investigators understand the similarities and differences between program evaluation and research.

- When program activities respond to a research question or a hypothesis, and the information collected contributes to generalizable knowledge, then the program includes a research component (i.e., beyond the context of the specific institution[s] conducting the evaluation).
- The IRB determines whether these projects are research on a case-by-case basis.
- The IRB makes this determination by evaluating a group of factors, including the purpose and intention of the project, level of risk, and methodology.
- Publishing or presenting program evaluation findings does not automatically mean the project is research.

For program activities to fall within the category of research, the IRB assesses whether these activities meet the definition of research and whether the project involves human subjects. With the understanding that mentoring programs involve humans, we focus on the definition of research.

According to the federal regulation¹, research per human subject protection regulations means a systematic investigation (including research development, testing, and evaluation) designed to develop or contribute to generalizable knowledge.

The keywords in the definition of research are *systematic investigation* and *generalizable knowledge*. The dictionary defines systematic as having a method or plan, possibly concerned with classification. Definitions of investigation include a detailed or careful examination, exploration, or learning of the facts about something complex or hidden. Attempting to answer a question or prove/disprove a hypothesis indicates that an activity is a systematic investigation² Contributing to generalizable knowledge means that there is intent on sharing information about the mentoring program with others.

The IRB evaluates several factors in determining if program activities fall within evaluation or research. Both research and evaluation activities can share some of the same outputs, but with different content. For example, when designing the program's activities, program coordinators may choose to publish findings from the program. While both evaluation and research findings may be published, the content would differ. A publication stemming from the evaluation would describe the results of the evaluation. In contrast, a publication from the research activities might describe the effects of the mentoring program. To clarify this, we provide Table 14.1. This table is condensed and modified to focus on mentoring. In the first column are eight common elements that help distinguish between program evaluation and research. The first common element is intent, and the last is the dissemination of results. The second and third columns explain the difference between evaluation and research for each of the eight common elements.

Table 14.1*Common Elements of Evaluation versus Research for Mentoring Programs*

| | | |
|--|--|--|
| Intent | The intent is to evaluate a specific academic mentoring program and only provide information for and about that particular program. | The intent is to do a systematic investigation, including research development, testing, and evaluation designed to contribute to generalizable knowledge. The data will be used to draw conclusions for the larger academic mentoring field. |
| Focus | The focus is on the mentoring processes, products, or programs. | The focus is on the mentoring population (human subjects) or strategies the mentors utilize. |
| Subject population | Statistical justification is not used to determine the sample size. | Statistical justification or other disciplinarily appropriate methodology determines the sample size. |
| Design and desired outcome | <p>The mentoring program is designed to assess the effectiveness of or improve a process, product, or program via:</p> <ul style="list-style-type: none"> · needs assessment · process, outcome, or impact evaluation · cost-benefit or cost-effectiveness analyses <p>May involve a comparison of variations in the mentoring program.</p> | <p>The mentoring program or subsequent inquiry is designed to answer a question or test a hypothesis to develop or contribute to the scientific storehouse of knowledge or theory within the mentoring field via:</p> <ul style="list-style-type: none"> · procedures, component(s), or analyses (i.e., involving combining data with other projects); · randomization of individuals to different processes or interventions; · novel research ideas, experimental activities that are not yet known to be efficacious; or · expanded sites or literature reviews. <p>May be designed to be descriptive or prove a relationship, correlational, or causation.</p> |
| Effect on standard procedures or normal activities | The evaluation of the mentoring program rarely alters the standard procedures while the mentoring project is ongoing. | An experiment or nonstandard intervention may alter the mentoring program's standard procedures or normal activities. |
| Funding | The mentoring program may be unfunded, funded by the university, or externally funded by an agency focused on mentoring programmatic activities. | The mentoring program may be unfunded, funded by the university, or externally funded by an agency focused on mentoring research. |
| Effect on program or practice evaluated | Findings of the evaluation are expected to directly affect the conduct of the program and identify improvements. | Findings of the study are not expected to directly or immediately affect the program, although they may also be used for this purpose. |
| Dissemination of results | The results of the program evaluation may be published. The intention is to disseminate details of the program's effectiveness and not contribute to generalizable knowledge. | The desire to share the effects of the mentoring program impacts the choice of procedures, design, and analyses to strengthen generalizability and extend the program's findings. |

Note. Adapted from “Program Evaluation vs. Research: Do I Need to Submit for an Exemption or IRB Approval?” by Julie M. Eiserman, August 23, 2023, p. 3, Office of Intramural Research ([https://irbo.nih.gov/confluence/download/attachments/70321066/Program Evaluation vs. Research.pdf?version=1&modificationDate=1545630790161&api=v2](https://irbo.nih.gov/confluence/download/attachments/70321066/Program%20Evaluation%20vs.%20Research.pdf?version=1&modificationDate=1545630790161&api=v2)).

After reviewing Table 14.1, it can be helpful for program coordinators to answer the following questions to help determine if program activities fall within the categories of evaluation or research.

- Is the *intent* to systematically test hypotheses and draw conclusions for the larger academic mentoring field? If yes, the activity is likely research.
- Is the *focus* on human subjects? For example, strategies mentors might utilize. If yes, the activity is likely research.
- Is statistical justification used to determine the methodologically appropriate sample size of the *subject population*? If yes, the activity is likely research.
- Does the *program’s design* contribute to the scientific storehouse of knowledge by using comparison groups? If yes, the activity is likely research.
- Does the *program’s design* contribute to the scientific storehouse of knowledge by including novel activities in the mentoring field? If yes, the activity is likely research.
- Is the program *designed* to assess relationships, correlations, or causations among variables of interest that contribute to the scientific storehouse of mentoring knowledge? If yes, the activity is likely research.
- Are the *program’s standard* procedures altered by an experiment or nonstandard intervention? If yes, the activity is likely research.
- Is the program *funded* by an agency focused on mentoring research? If yes, the activity is likely research.
- Do the *dissemination* goals impact the program’s procedures, design, and analyses to strengthen generalizability and extend the program’s findings? If yes, the activity is likely research.

Suppose the answer to any of the above questions is yes. This suggests that the program’s activities involve human subjects and contribute to generalizable knowledge, falling within the research category. The program coordinator should work with their respective IRB to make this determination and ensure the research is conducted according to IRB standards. The second section of this chapter guides the program coordinator as they navigate their respective IRB.

Navigating the Institutional Review Board

Research with human participants has been a way of acquiring new knowledge since time immemorial. While this tradition has a rich history of gainful, ethical scientific inquiry, there is also a darker side to scientific exploration using human research subjects. Institutional review boards (often called research ethics boards, human research ethics boards, or research review boards in contexts outside of the United States) were developed in response to a fraught history between those conducting

the research and those being researched. Infamous examples abound, such as the Tuskegee syphilis study, the Stanford prison experiment, and more recent ethical blunders like the Facebook contagion study and experimental contributions to the Linux kernel by researchers at the University of Minnesota.

In short, IRBs exist to ensure that research with living people as subjects is conducted according to certain ethical principles. Those principles depend on the context of the research work. In the United States, IRBs adhere to the *Ethical Principles and Guidelines for the Protection of Human Subjects of Research* document created by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. This document is called the *Belmont Report* (so named for the location where the commission met and created this guide in 1979: the Smithsonian Institution's Belmont Conference Center).

You will find an IRB at any institution or organization that accepts federal funding for research with human subjects, as defined in the previous section of this chapter, and even within some institutions or organizations that do not receive funding but wish to ensure that their human subjects research portfolio is being conducted ethically. If an institution does not have its own IRB, it likely contracts with another institution's IRB or an independent IRB for its reviews. Typically housed within an office for research or an office for academic affairs, IRBs are structured to meet the requirements of 45 CFR 46: a chair, a nonscientist member, a scientist member, a person unaffiliated with the institution, and a prisoner representative, if that organization conducts research with incarcerated individuals.

Policies and procedures at each organization define who can determine whether a project is considered research with what is considered human subjects. Most institutions leave the final say to their IRB. Thus, it is prudent to check with your IRB about whether your program activities might constitute research with human subjects, or HSR (human subjects research). Most institutions make a straightforward and fast process available wherein a researcher can obtain official documentation from their IRB that the process is not HSR—typically called an NHSR determination or a request for determination process. Where it is desirable, many institutions' IRB staff will even help faculty reshape their project so that the project falls outside the IRB's jurisdiction. Successfully navigating the IRB begins with understanding what the IRB expects from you.

What Does Your IRB Expect From You?

Each IRB approaches its work in different and distinct ways. As you build your evaluation plan within your team, consider using your institution's IRB application template as a guide. Every IRB application is different, but you can expect to address the following items no matter where you are submitting:

- purpose of the project
- research staff who will contribute to the research program's effort
- whom will the participants be
- how will they be asked to participate
- the procedures for your evaluation activities and research activities
- risks and benefits of the project

- how the research team will manage/protect privacy and confidentiality
- whether and how informed consent will be obtained

Before you ever start to fill in an application (called a *protocol* in the IRB world), if you have thought through and addressed these bulleted items as a team, you will be in excellent shape for submitting your application to your IRB. Once you have outlined your proposal for your evaluation and research efforts, we suggest meeting with someone from your institution's IRB. Request a consultation with the staff who review these submissions and ask them to talk with you about how to make the review process go smoothly. Based on our experience at Utah State University, our IRB office team finds that most common pitfalls in the review and approval process for mentorship projects stem from a lack of clarity between the activities of the program, the activities of evaluation, and the activities of research. As you consult with your IRB staff, make sure they understand these distinctions. If there are new, extra, or experimental things that you are doing to create generalizable knowledge, share these with your IRB consultant. A good IRB staffer will pick up on the nuances between program evaluation and program research and help you shape a protocol that can be reviewed swiftly without too many clarifying questions.

Review Processes and Timelines

The review process you will undergo depends very much on the structure of your evaluation and/ or research efforts. A scenario in which you are collecting readily available information from/about participants, not using comparison groups, working with adults, and evaluating the program at the request of your institution will look very different from a scenario in which you are sorting participants into comparison groups, conducting additional surveys to address specific research questions, or working with children. Both are valid ways to approach evaluation or research efforts, and you should choose the approach that best fits the purpose of your program. These distinctions will help you understand the review and process of the IRB.

Under the first scenario mentioned, you can expect your IRB to declare your research *exempt*. Exempt, in this context, does not mean that you do not have to submit to your IRB; instead, your project will be exempted from most requirements outlined in 45 CFR 46. Note that this would not exempt the project from other requirements, such as those enshrined in the Family Educational Rights and Privacy Act (FERPA) regarding the protection of identifiable student records, as well as any requirements your institution might put in place regarding issues from allowable data access/storage to what email accounts are permitted to be used. Usually, your IRB will inform you of these requirements either in their review process or by using ancillary review processes, which bring in other offices to ensure that the institution has prepared you for a successful and compliant project.

Under the second scenario outlined above, researchers should plan for at least an expedited review process, meaning that it is eligible for a process that is faster than a review by the full board (sometimes called convened IRB review), using a process outlined by the policies and procedures governing the IRB. In some cases, depending on the design of the intervention groups, the review might occur via the full board. Timelines for these processes vary. IRBs accredited by the Association for the Accreditation of Human Research Protection Programs (AAHRPP) will compile, at least annually, timelines for their review processes. Therefore, asking how long a review process will take is perfectly acceptable. The most recent timelines from AAHRPP show that researchers should expect a turnaround time of 20–50 calendar days on their applications.

Researcher Responsibilities Post-Approval

Once the protocol is approved, it would be surprising if everything went exactly as you planned in your initial protocol. The realities of research implementation, working with humans, and the unexpected necessarily mean that changes will occur. Your IRB expects this to be the case and will have policies and procedures outlining when a researcher must file an amendment (or modification) to their protocol. Before implementing the change with your participants, revisions must occur, so plan for adequate lead time! If a change needs to be made before the IRB reviews and approves it, it becomes a reportable event that you must disclose to the IRB. In all cases, researchers must keep a line of communication open with their IRB regarding their project.

Another vital aspect to consider post-approval is informed consent. The process and procedures of obtaining informed consent will be clarified during the review. However, that does not mark the end of the researchers' responsibility to keep participants informed. As changes occur, it is crucial to ensure that your participants remain up-to-date on essential facets of the project. Informed consent is an ongoing process, not a one-time interaction, and researchers remain responsible throughout the project's lifetime for adequately informing their research participants.

Post-approval monitoring is another way researchers might interface with their IRB once initial approval has been obtained. That might be as simple as an email check-in with the research team or as involved as a full audit of study records, depending on your IRB's policies and procedures. In all cases, it is critically important to respond promptly and honestly. Some other important considerations for the life cycle of the protocol include:

- keeping research staff up-to-date
- ensuring all members of the study team maintain adequate training
- knowing where to find necessary documentation, such as approval letters, continuation review letters, amendment approvals, and disapprovals, and any status reports you previously submitted to the IRB
- updating the IRB on the status of your project

Closing a Protocol

Generally speaking, it is time to close your protocol when the interventions with participants are done, analyses complete, and identifying information about participants destroyed. IRBs will have a process for doing this; in some systems, it is as simple as pushing a button in your protocol-management system. A full audit might be appropriate before the project wraps up. Be sure to keep informed-consent documentation for at least 3 years following the closure of the protocol and finalize compliance with any data-management plans you might have submitted previously. If you commit to sharing findings with research participants, ensure this is completed before closing your protocol.

The most important thing to remember is that your IRB supports your research while prioritizing your research participants' perspectives, rights, and welfare. IRB staff and board members are not trying to overregulate, protect the institution's interests, or find wrongdoing, so keeping open lines of

communication and periodically checking in will benefit all parties during the life cycle of your mentorship evaluation project.

The first two sections of this chapter prepare the program coordinator to distinguish between program evaluation and program research and how to successfully navigate the IRB when their mentoring program does include research. The third section of this chapter focuses on how creating a theoretically and methodologically sound program enhances the program's research activities.

Creating a Theoretically and Methodologically Sound Mentoring Program

Chapter 2 of this book describes the essential role of theoretical frameworks in designing mentoring programs. Chapter 13 provides frameworks to guide assessment and evaluation efforts. Finally, this section of Chapter 14 discusses theoretical frameworks and program methodology pertaining to research. The theoretical framework and methodological principles discussed in this section will improve the program research and overall quality. In reviews of the scientific literature on mentoring programs in higher education (Crisp & Cruz, 2009; Gershenfeld, 2014; Jacobi, 1991; Tinoco-Geraldo et al., 2020), scholars in the field have suggested that research into formal mentoring programs lack much of the theoretical and methodological rigor that is common in other areas. This section guides the reader in strengthening their program's research by connecting it to a sound theoretical framework and rigorous methodology.

Theoretical Frameworks

We continue this discussion on theoretical frameworks in the context of research because methodologically sound research benefits from a solid theoretical foundation. Before proceeding with this section, we encourage readers to skim the four case studies in Chapters 16 through 19, where the mentees are undergraduate students; the two case studies in Chapters 20 and 21, where the mentees are graduate students; the three case studies in Chapters 22 through 24, where the mentees are faculty members; the two case studies in Chapters 25 and 26, where university staff are the mentees; and the one case study in Chapter 28, on networked mentoring.

Suppose you are a program coordinator in your college, and an associate dean asks you to address the high attrition rate of undergraduate students. As you familiarize yourself with the literature on student attrition, you come across Vincent Tinto's (1993) landmark book *Leaving College: Rethinking the Causes and Cures of Student Attrition*. On page 147, you read, "Effective retention programs are committed to the development of supportive social and educational communities in which all students are integrated as competent members." As you reflect on Tinto's comment, you begin to appreciate that a faculty-to-student mentoring program could help develop these supportive social and educational communities, which could lead to higher retention. These relationships are summarized as follows:

Faculty-to-Student Mentoring Program → Student Retention

As you continue to explore the literature on theoretical frameworks for formal mentoring programs, you ask yourself, How does a mentoring program lead to student retention? Through the continued reading of Tinto's social integration theory, key constructs such as a *sense of belonging* and *student*

retention crystalize, and relationships between these constructs begin to take shape. As you continue your exploration of theoretical models, you also are drawn to Kram's *mentor functions* (Kram, 1985) and social learning theory (Bandura, 1977). As you reflect on key theoretical constructs, you start to make connections such as a mentor who provides academic subject knowledge, career guidance, and psychosocial support will become a role model for the mentee. By providing these services, the student feels like they belong to the university family, which will increase retention rates.

Faculty-to-Student Mentoring Program → Sense of Belonging → Student Retention

Developing key constructs that are theory-driven and clearly stating the relationships between these constructs starts to provide a model for how you think your intervention will impact the mentees of your program. This theory-of-change model is essential to developing an effective program and program research. In this chapter, we use the theory of change to connect key constructs from theoretical models to the program's desired outcomes. Chapter 13 also uses the theory of change to describe logic models that explicitly connect resources, activities, outputs, outcomes, and impacts (see Figure 13.1).

Theory of Change

As you continue the exercise above, asking how and why your mentoring program is supposed to lead to your desired outcomes, you will develop a theory of change that can be summarized with a series of if/then statements. First, you need to create a diagram of your theory of change, which can guide you in implementing your program's research. You can find an example of an effective theory-of-change diagram in Appendix A of Chapter 18. An abbreviated³ version of the if/then statements from that case study are as follows:

- IF mentees **enroll in the mentoring program**, THEN the mentor will provide **academic expertise, career guidance, psychosocial support, and role modeling**.
- IF mentors provide mentees with academic expertise, career guidance, psychosocial support, and role modeling, THEN mentees will **successfully adjust to the university and feel like they belong there**.
- IF mentors help mentees successfully adjust to the university and gain a sense of belonging, THEN mentees will **connect to an academic discipline and develop goals and a plan to achieve them**.
- IF mentors help mentees develop a plan to achieve their goals, THEN mentees will **increase their persistence, retention, grade point average, and graduation rates**.

Describing the theoretical links between mentoring and student retention is not just an intellectual exercise; it shifts the focus of what is emphasized. With a theoretical framework, links between mentoring and the dependent variables being researched can be explained. Jacobi (1991) cautioned that mentoring programs might be inadequately developed when models or frameworks of mentoring remain implicant and lack clarity. In summary, to reach the intended outcomes of increased persistence, retention, grade point average, and graduation rates, the mentors in this program will need to provide academic expertise, career guidance, psychosocial support, and role modeling to the mentees. Spending time developing a clear and logical theory-of-change model offers additional

benefits as it guides the creation of an operational definition and clarifies research processes.

Clear Definition of Mentoring

When developing a theory of change, it is essential to begin with a clear definition of mentoring. A lack of a clear conceptual definition is problematic because it limits the ability to measure what constitutes a successful mentoring experience. Furthermore, a lack of clarity about what is being measured also contributed to weak research designs commonly found in the mentoring literature (Crisp & Cruz, 2009; Jacobi, 1991). Lastly, when key constructs are not made clear, it is difficult to replicate the program and program research, hindering the advancement of the science of mentoring.

In Chapter 1 of this book, Garvey describes the challenges of creating a singular definition of mentoring and instead advocates for a straightforward process that program coordinators can follow to develop their definition unique to their context. In addition to Garvey's work in Chapter 1, we find the work of Dominguez (2012) and Dominguez and Kochan (2020) helpful in guiding program coordinators' efforts in developing an operational definition for their mentoring program. Dominguez (2012) analyzed over 457 definitions of mentoring and found one overarching dimension and five elements commonly repeated. The overarching dimension is that mentoring is first and foremost a developmental relationship. The five elements included in most definitions were: (a) qualifier defining the desired qualities of the relationship; (b) defining word(s) specifying the type of relationship; (c) participants providing and receiving mentoring; (d) functions or activities in which participants engage to achieve desired outcomes; and (e) outcomes or achievements the mentor and mentee expect to accomplish. As program coordinators develop the operational definition for their program, they should let this process be informed by the theoretical frameworks used in the program. The purpose of this section is to advocate that operational definitions should be connected to the theoretical frameworks being used. When these connections are apparent, they clarify which constructs will be used and how they will be defined.

To help the reader make these connections, we again highlight the case study in Chapter 18. In this case study, there were three theoretical frameworks used: (a) Kram's mentor functions (Kram, 1985), (b) social learning theory (Bandura, 1977), and (c) social integration theory (Tinto, 1987, 1993). Additionally, the work of Nora and Crisp (2007) and McWilliams (2017) influenced the development of a clear definition of mentoring. New program coordinators may erroneously think they should only choose one theory to base their mentoring program on. However, Gershenfeld (2014) suggests that modern mentoring programs should use multiple guiding theories. Based on these three theories, the emerging constructs of interest were academic expertise, career guidance, psychosocial support, role modeling, successful adjustment to the university, and a sense of belonging or connectedness. With these constructs in mind, Spears, Hales, and Lewis, authors of Chapter 18, developed the following definition of mentoring. We have highlighted how four of Dominguez and Kochan's (2020) five elements factor into this definition. The qualifier element did not cleanly fit into the following definition:

Building a purposeful and personal relationship (defining word) in which a more experienced person (mentor)(participant) provides guidance, feedback, and support (functions or activities) to facilitate the growth and development (outcome) of a less experienced person (mentee) (participant). Operationally,

mentors provide mentees with services such as (functions or activities):

1. *Academic subject knowledge and institutional support*
2. *Education/career exploration and goal setting*
3. *Psychosocial support*
4. *Role modeling.* (Chapter 18)

Examining this definition, we hope the reader can see how this clear definition of mentoring influenced the overall development of the theory of change illustrated in this case study. Now that we have discussed how theory affects the development of the mentoring program, we now turn to the main focus of this section, which is to understand how theory connects to research in mentoring programs.

Theoretical Framework as a Guide for Research

In empirical (research) studies, theory guides a researcher in understanding what is important to measure as part of the research project. Considering the theory of change presented above, properly researching this program will include measuring many variables, including whether the mentor provides academic expertise, career guidance, psychosocial support, and role modeling. We would also want to measure the mentee's adjustment to the university, sense of belonging, academic goals, motivation, and our intended outcomes of persistence, retention, grade point average, and graduation. As noted in their review, Tinoco-Giraldo and colleagues (2020) found that more studies on mentoring programs in higher education have identified a theoretical foundation than was present in previous reviews (Jacobi, 1991; Crisp & Cruz, 2009). However, although more studies identified a theoretical foundation, only some linked theory with methodology. Most studies measured satisfaction with the mentoring relationship and called that enough; however, we need other elements to understand effective mentoring. The most refined theoretical models, such as Kram's mentor functions (Kram, 1985), Hunt and Michael's (1983) model of mentoring, O'Neil and Wrightsman's (2001) sources of variance theory, and Tinto's (1993) social integration theory, have rarely been effectively researched (Johnson et al., 2010).

Beginning with a firm theoretical foundation helps develop a mentoring program and sets up an effective research program. In the next section, we introduce the reader to basic research methodology and how it impacts the research findings.

Methodological Rigor

With an understanding of how critical a theoretical framework is, we can now discuss sound methodological principles and how these lead to an effective program. We recommend that coordinators of mentoring programs in higher education audit a course on research methodology. Such classes can be commonly found in psychology, sociology, and other related departments. These courses will provide a more in-depth look at research methodology, whereas this section is meant to be a primer on the topic. The following information will increase the validity of your research program's findings. When conducting research, it is vital to recognize and address internal and external validity threats.

Threats to Internal Validity

Internal validity refers to the extent to which your findings can be trusted. For example, suppose in the research of your mentoring program, you conclude that retention increased among your participants. This result may be because of your program. However, it could also be due to some unrelated factor. Sound methodological design will help reduce the threat to the internal validity of your findings so you can be confident in your results.

Research Design. When testing the effectiveness of a mentoring program, a randomized controlled trial (RCT) is considered the gold standard (Webber & Prouse, 2018; if you are interested, you can find a critique of RCTs in Grossman and Mackenzie [2005]). In an RCT, participants are randomly selected from a population and then randomly put into a treatment group or a control group. The treatment group receives the program procedures, while the control group does not. After the duration of the program, researchers compare the two groups on the outcome variables. If we were to use an RTC design with a mentoring program, it would involve randomly selecting students from the entire student body. Half of those students would be chosen randomly to participate in the mentoring program. In contrast, the other half would not participate and instead would be used as a comparison at the end of a specified period, for example, one academic year.

While it is generally considered the best approach for studying programs like these, RTC might not be feasible to research mentoring programs. Because there is support for the positive effects of academic mentoring (Eby et al., 2008; see also Chapter 4 in this collection), we believe it unethical to employ a classical research design such as an RTC with random assignment to the treatment group and control group, thus denying the control group access to the mentoring program. In addition, it would be problematic to deny the program to individuals seeking the additional support associated with a mentoring program and instead randomly choose which students from the student body would be eligible for the program.

A comparison group is still necessary to address threats to the internal validity of the evaluation. For example, imagine that you evaluate your mentoring program and find that students who participated increased their GPA from the previous year. If you had included a comparison group, you might have also found a similar increase in that group. On the other hand, perhaps a year's experience was enough to improve your GPA, and your mentoring program did nothing. Therefore, when considering a research design for mentoring in academia, we recommend using either a waitlist control group or a quasi-experimental propensity-matched control group for comparison purposes.

In the waitlist control group design (see also the "Control Group Comparisons" section of Chapter 13), a group of potential mentees who do not receive the mentoring is put on a waiting list to receive the mentoring intervention after the treatment group receives the intervention. Conceptually similar to a waitlist control group design, a delayed-start design could easily be applied to mentoring programs in higher education. For example, suppose you are tasked with creating a university-wide mentoring program for faculty of color regarding feelings of isolation, lack of representation, and suboptimal retention, as described in Chapter 23. Using a delayed-start design, you could implement the program for one college in year one, another in year two, and so on. This delayed-start design will provide a naturally occurring control group for comparison purposes.

In a quasi-experimental propensity-match control group, the control group consists of matched individuals like the participants in the treatment group. These matches are made from variables of interest, such as GPA, race, first-generation student status, or other demographic variables. This example is the same control group comparison explained in the “Control Group Comparisons” section of Chapter 13. Laura Lunsford, the author of Chapter 13, states that institutional research offices can provide the identity of people similar to the treatment group for comparison purposes. In addition, the “Outcomes of the Program” section of Chapter 18 describes persistence rate comparisons between a treatment group and a propensity-matched control group.

Time Points for Data Collection. Jacobi (1991) found that most empirical research on mentoring relied on retrospective, correlational designs using small samples with data collected at a single time. All of these present a threat to internal validity. Even with a comparison group, it can be challenging to determine if a change has occurred without multiple measurement points. If you gather data on the same participants over some time, this is called a longitudinal study. If you collect data one time on a sample of a population, this is called a cross-sectional study. Crisp and Cruz (2009), Gershenfeld (2014), and Jacobi (1991) stress the need to collect data at multiple time points. Jacobi (1991) further suggests that collecting data at multiple time points is important because it is yet to be determined how long it takes for mentoring effects to emerge.

Additionally, if you reflect on the theory of change presented earlier, it is clear that multiple measurement time points are necessary to test such a model. The process of receiving support from a mentor, feeling connected to the university, developing and then working toward academic goals, and finally accomplishing those goals is unlikely to occur in a single semester or academic year. Therefore, we echo the recommendation to include more than one measurement point in researching mentoring programs.

Mentoring programs in higher education have natural times to collect data, such as the beginning of an academic year, the end of the fall semester, and the end of the winter or spring semester. For example, suppose you are a university staff member and desire to create a mentoring program to empower staff members, similar to the case study in Chapter 26. Universities must be fully staffed at the beginning of an academic year, so the beginning of the fall semester provides an opportune time to collect data on new staff employees. Data can be collected at the end of the fall and spring semesters to gauge staff members’ sense of belonging.

Clear Identity of Variables. Identifying the variables is essential for two reasons. First, it helps other researchers replicate future studies using the same constructs, dimensions, indicators, and attributes. Second, and more important, clearly identifying the variables and discussing their connection to the theoretical framework and operational definition make it explicit how the independent and intervening variables are expected to influence the dependent variables.

Threats to External Validity

External validity refers to how well your findings can be generalized to other populations. For example, would you have similar outcomes if you took any case studies in this book and modeled a similar program at your institution? Crisp and Cruz (2009) recommend that potentially extraneous

variables such as institution type, mentee and mentor attitudes, and other characteristics of mentee and mentor—for instance, gender or ethnicity—might affect the external validity of findings. Gershenfeld (2014) points out additional threats to external validity, such as small sample sizes, single geographical locations, and narrowly focused programs. We will begin with a discussion of Gershenfeld’s critiques, which focus on samples, and then conclude with a discussion of Crisp and Cruz’s comments on variables.

Sampling. Scientists use samples when researching because it is likely impossible to gather data from the entirety of the population of interest. If it were possible, it would represent an enormous financial and logistical burden. Thankfully, a sample from the population can be considered representative of the larger population thanks to the central limit theorem (a statistical principle beyond this chapter’s scope). However, a poorly designed sample will limit the findings’ external validity.

One thing that is important to consider is the size of the sample. Gershenfeld (2014) notes that more than a small sample size might be needed to find the mentoring program’s effect on its participants. We call this a Type I error, when there is an effect but our study has failed to find it (false negative). The solution to this problem is gathering a manageable sample. That might make your statistical tests too sensitive where you commit a Type II error—when there is no effect but your study has found one (false positive). Too small and too large of samples both venture into the realm of unethical because they waste the participant’s time and effort for an unscientific outcome (for further reading on this topic, we recommend Martinez-Mesa et al. [2014]).

Another of Gershenfeld’s (2014) critiques is when programs use a single geographical location. When feasible, academic institutions should establish the mentoring program at multiple sites. If the academic institution has multiple campuses, this can easily be obtained. If the institution does not have numerous campuses, then program coordinators should strive for implementation beyond one site. For example, suppose that a specific graduate program within the college of education proposed implementing a peer mentoring program in which advanced graduate students mentor new incoming graduate students. To improve external validity, the program coordinator could propose that this program be offered to all incoming graduate students within the college.

Narrowly Focused Program. Gershenfeld’s (2014) final critique is that some programs are too narrowly focused. A program like the one described above, with peer-to-peer mentoring among graduate students, may have limited generalizability to other programs, such as one designed for faculty of color. The more general the mentoring program is, the more general its sample will be, contributing to greater external validity. However, not all mentoring programs are going to be general.

Often a mentoring program is designed to address a need of a specific population. In these cases, a solid theoretical foundation, a clear definition of mentoring, and clearly defined and psychometrically sound variables will improve the program’s generalizability.

Extraneous Variables. When evaluating a mentoring program, it would be easiest to include only the primary variables of interest in your study. Methodologists call these the independent and dependent variables. The dependent variables are the outcomes you are interested in, and the

independent variables are thought to be associated with these outcomes of interest. As you develop a theory of change, you will find that many variables will be important to study. In Appendix A of Chapter 18, you can see that there is an associated assessment to measure each of the constructs identified by the program coordinators in their theory of change. Sound methodological design will also include additional variables, like those mentioned by Crisp and Cruz (2009): institution type, mentee and mentor attitudes, and other characteristics of mentee and mentor—for instance, gender or ethnicity. While your theory of change might be sound, it is possible that any positive effects found could be the result of one of these extraneous variables. For instance, suppose you fail to include gender as a variable in your evaluation. You find that the program overall is effective; however, if you had included gender as a variable, you might have seen a significant improvement for female mentees with female mentors and little to no effect for any other group.

Institution Type. Gershenfeld’s (2014) recommendation for methodological rigor requires identifying the type of institution performing the research. For example, is the institution a community college or a four-year university? Do faculty at the institution have as their primary role teaching or research? Is the institution primarily a residential campus or a commuter campus? Is the institution located in one city or are satellite campuses spread throughout the state? Students attending these different types of institutions will have some baseline differences, and not disclosing that information creates a substantial threat to external validity. However, as mentioned above about narrowly focused programs, a robust theoretical foundation will help to minimize this threat to external validity.

Mentee and Mentor Attitudes. Program coordinators should gather attitudinal information to see how it impacts the mentoring program’s outcomes. Examples of attitudinal information could be satisfaction with the mentoring relationship, perceived effectiveness of the mentoring program, satisfaction with the mentoring program, and mentoring program understanding. It might be that the level of understanding a mentee has of the program itself —such as the procedures or what is expected of the mentee and mentor —is associated with the mentee’s persistence through the program. Remember that research aims to increase the general knowledge of the topic, and nuances like this would be valuable to other programs.

Other internal attitudes, such as motivation to participate in the program, could also impact desired outcomes. Motivation is crucial to persistence (Ryan & Deci, 2000) and is especially so in higher education (Müller, 2008; Simon et al., 2015).

Characteristics of Mentors and Mentees. The last extraneous variable identified by Crisp and Cruz (2009) and supported by Tinoco-Giralso et al. (2020) was the characteristics of mentors and mentees. Most programs will gather data on demographic characteristics such as gender, age, and race. However, if the needs assessment described in Chapter 5 identifies other factors critical to the study, these characteristics should be collected. For example, as a program coordinator, you were tasked to develop a faculty-to-student mentoring program to address student attrition. As part of the needs assessment, you discovered that the most vulnerable students for not returning to the university were students who had not picked a major or had an undeclared major. If this was the case, this characteristic should be gathered and assessed.

A final note to our discussion of research methodology is that Tinoco-Giralso et al. (2020) also

recommended that measurements used to assess the mentoring relationship quality be validated. Using valid measures is essential for all of the variables under investigation. The last section of this chapter addresses the issue of valid measurement.

Measurements for Academic Mentoring Programs

Assessment is integral to any research and scientific endeavor to improve a project's quality and outcomes while gaining insight into the question(s) under examination. Unfortunately, assessments supporting mentoring programs have long suffered the same inconsistencies as the definition of mentoring faces, often lacking agreement on the essential functions of the relationship and criteria for evaluating its effectiveness (Berk et al., 2005). Noe (1988) indicated that mentoring lacks quantitative measures for the functions mentors provide to their mentees in the assessment field. There are commercially available assessments, especially for career mentoring programs; however, a disconnect exists between research-based mentoring scales and the instruments that practitioners use in several of these products (Gilbreath et al., 2008). Many of the tools available are designed to evaluate specific programs only, measuring the value of the mentoring functions or the frequency of mentoring. Jacobi (1991), Crisp and Cruz (2009), and Gershenfeld (2014) have all indicated that programs lack rigorous and valid instruments to measure their intended effect and outcomes.

Existing Tested Constructs for Program Assessment

As mentioned earlier in this chapter, the work of Scandura (1992), Noe (1988), Ragins and McFarlin (1990), Allen and Eby (2003), Allen et al. (2006), Hurtado et al. (2007), Ragins and Scandura (1999), and Crisp (2009) give a foundation for mentoring practitioners and researchers to use psychometrically sound assessments for both mentors and mentees. Their collective work provides assessment items supporting several constructs, including psychological and emotional support, degree and career support, academic subject knowledge support, the existence of a role model, satisfaction with the mentoring relationship, perceived program effectiveness through benefits for mentors, psychosocial support, sense of belonging, and success at managing the academic environment. These constructs, their associated research, and descriptions of the instruments are reviewed next. When different studies explore similar constructs, they are grouped together.

Crisp (2009) provides research on the constructs of *psychological and emotional support, degree and career support, academic subject knowledge support, and the existence of a role model* by examining the validity of the college student mentoring scales (CSMS) (Crisp, 2009). This instrument uses eight items to measure psychological and emotional support, six items to assess degree and career support, five items pertaining to academic subject knowledge support, and six items for the existence of a role model. All 25 items in Crisp's CSMA use a five-point Likert-type scale of *strongly disagree to strongly agree*.

Allen and Eby (2003) explore the construct of *satisfaction with the mentoring relationship through mentor effectiveness* by focusing on relationship learning and quality. They used two mentorship quality items developed earlier by Noe (1988) and Ensher and Murphy (1997). Their survey collected demographic information such as age, race, gender, education, institutional longevity, and occupation from professional mentors. It includes five items related to relationship learning and five items measuring relationship quality using a five-point Likert-type scale of *strongly disagree to strongly agree*.

Two constructs, *satisfaction with the mentoring program* and *perceived program effectiveness through program understanding and training*, are supported by the work of Allen et al. (2006). Their 17-question survey included seven items on perceived program effectiveness, four items on mentor commitment, four items on program understanding, and two items on program characteristics. A five-point Likert-type scale is used for all items, with the exception of the two for program characteristics. These included a yes/no and a four-indicator response for how much input mentors/protégés had on whom they were matched with.

Benefits for mentors come from Ragins and Scandura (1999), who examined the costs and benefits of being a mentor specifically for executives in a nonformalized mentoring setting. Their instrument uses a seven-point Likert-type scale ranging from *strongly disagree* to *strongly agree*, with 17 cost items and 24 benefit items. Five factors emerged for the cost items: more trouble than it is worth, dysfunctional relationship, nepotism, bad reflection, and energy drain. Five factors also appeared for the 24 benefit items: rewarding experiences, improved job performance, loyal support base, recognition by others, and generativity.

Psychosocial support benefits of participating in a mentoring program come from multiple studies. Dreher and Ash (1990) developed an 18-item instrument using a five-point Likert-type scale from *not at all* to *to a very large extent* by examining a number of the career and psychosocial functions that Kram identified in 1985. Their work considered mentoring relationships in a professional environment, surveying business program alums from two universities. Tenenbaum et al. (2001), building off the work of Dreher and Ash, developed a five-part survey measuring the satisfaction of graduate students' advisor–advisee relationships. Part one of this instrument includes 19 items measuring three factors: psychosocial, instrumental, and networking support of their graduate advisors.

Sense of belonging and *success in managing the academic environment* constructs comes from Hurtado et al. (2007). Their instrument's questions were developed to investigate critical factors impacting first-year college transition for underrepresented minority students in biomedical and behavioral sciences programs. The instrument's questions can be part of the ongoing monitoring of students' transition experiences and as part of a university's climate studies (Hurtado et al., 2007). Their five-item sense of belonging construct uses a three-point Likert-type scale from *unsuccessful* to *completely successful*. Successfully managing the academic environment construct is three items using a four-point scale of *strongly disagree* to *strongly agree*.

Existing Assessments Supporting Higher Education Mentoring Programs

In addition to the instruments mentioned above, which allow programs to evaluate certain construct areas, packaged evaluative tools are available, some of which are commercially available. A sampling of existing assessments and their descriptions are included in the following discussion.

Mentoring programs developed specifically for students participating in the medical field have several assessment tools available. The Mentorship Profile Questionnaire and Mentorship Effectiveness Scale were developed at Johns Hopkins University School of Nursing (Berk et al., 2005). The questionnaire contains four open-ended questions that allow mentees to describe their

relationship with their mentors and the outcomes of the relationship. The effectiveness scale consists of 12 items, assessing the relationship using a seven-point Likert-type scale. The Munich Evaluation of Mentoring Questionnaire (MEMeQ) is based on Berk's work and is designed for student mentees in the latter part of their medical training and examines the personal and content aspects of the mentoring relationship (Schäfer et al., 2015). The Mentoring Competency Assessment is a 26-item skills inventory evaluating the communication, expectations, understanding, diversity, independence, and professional development designed for clinical research mentors and mentees (Fleming et al., 2013). Finally, the Mentoring Evaluation Tool (MET) is a 13-item assessment instrument measuring the effectiveness of faculty mentors in one-to-one mentoring health science programs (Yukawa et al., 2020). The tool was developed at the University of California San Francisco's Schools of Dentistry, Medicine, Nursing, and Pharmacy. MET evaluates the effectiveness of mentors through five domains: meeting and communication, expectations and feedback, research support, career development, and psychosocial support using a seven-point scale from *strongly disagree* to *strongly agree*.

Previously discussed above in the construct section of this chapter, Crisp's (2009) College Student Mentoring Scale, a 25-item assessment measuring the psychological and emotional support, degree and career support, academic subject knowledge support, and role of a role model of the mentor by the mentee. In 2009, Gilbreath, Rose, and Dietrich assessed four commercially available mentoring assessments: the Allman Mentoring Activities Questionnaire (AMAQ), Mentoring in the Moment (MITM), Mentoring Skills Assessment (MSA), and Principles of Adults Mentoring Inventory (PAMI). PAMI is designed for career adults in academia mentoring adult learners. AMAQ, MITM, and MSA are for business settings; their findings showed that PAMI's content was valid, though they could not evaluate the instrument's reliability or validity of its construct criteria. However, PAMI may be helpful if mentors seek feedback to improve their practice and in training situations. The National Mentoring Resource Center provides a clearinghouse of handbooks, program manuals, and assessments. All of the assessment instruments featured by the National Mentoring Resource Center have a theoretical basis and have evidence of reliability and validity (National Mentoring Resource Center, 2016). Though their primary audience is youth mentoring programs, a handful of the assessments available are appropriate for mentees 18–25 years old. These include:

- Mentoring Processes Scale: A 26-item assessment using a seven-point Likert scale assessing mentor–mentee engagement designed for ages up to 21.
- Youth Strength of Relationship (YSoR) and Mentor Strength of Relationship (MSoR): A 10-item assessment for mentees and 14 items for mentors using a five-point scale, measuring both participants' experience perceptions of the mentoring relationship. They are designed for ages up to 21.
- Mentoring-Youth Alliance Scale (MYAS): A 10-item assessment using a four-point scale measuring the mentees' feelings regarding their mentoring experience. The MYAS is designed for ages up to 19.
- Problem-Solving Ability: A four-item, five-point scale assessment determining the mentee's problem-solving ability. This assessment is designed for ages up to 21.
- Career Exploration: A five-item assessment using a five-point scale to explore career fields. This assessment is designed for ages up to 25.

There are several published valid and reliable measurements to support mentoring programs in higher education. When determining what measurements to use for assessment, program evaluation, and program research, program managers and researchers must match measurements to their theory of change, with emphasis on the intended goals and outcomes of the mentoring program. Multiple measurements will need to be used to capture the nuances of the program's theory of change.

Conclusion

Chapter 14 uniquely contributes to this handbook by exploring the differences and similarities between program evaluation and program research. If choosing to do program research, this chapter guides the program coordinator as they navigate their university's institutional review board. Sound research methodology is enhanced when the theory of change is made explicit, connecting the theoretical framework, operational definition, and research methodology. Lastly, Chapter 14 provides examples of measurements that can be used for research, with some of these measurements also being appropriate for evaluation.

Lunsford, in Chapter 13, emphasizes that international standards for mentoring programs require assessment and evaluation as markers of an effective mentoring program. In Chapter 15, Castañeda-Kessel gives guidance for funding mentoring programs in academia. Some funding opportunities require mentoring programs to contain research and program evaluation. We conclude Chapter 14 by recommending to program coordinators and university leaders that their respective mentoring program includes research. One of the mentoring field's respected authors, Lillian Eby (2019), espoused conducting research in addition to the program's overall evaluation plan. In a workshop one of the authors attended, Eby trained program coordinators to include a research program. Eby's suggestions overlap with much of this chapter's content. Eby first advocates that program coordinators know the mentoring literature well enough to develop novel projects that advance the science of mentoring. Second, Eby advocates utilizing theory to inform evidence-based practices. Third, Eby explores how research design can be used systematically to test hypotheses and answer research questions. Fourth, Eby advocates for the use of psychometrically sound measures. Lastly, Eby described how to draw scientifically meaningful conclusions from the data.

After reading this chapter, we hope program coordinators and university leaders will consider adding a research component to their mentoring program. Adding a research component is not as daunting as it may seem. Coordinators are already carrying out many of the processes needed for research. Program coordinators can contribute to the science of mentoring with little additional effort by thoughtfully building a research program into their program's overall design.

References

- Allen, T. D., & Eby, L. T. (2003). Relationship effectiveness for mentors: Factors associated with learning and quality. *Journal of Management*, 29(4), 469–486. https://doi.org/10.1016/s0149-2063_03_00021-7
- Allen, T. D., Eby, L. T., & Lentz, E. (2006). Mentorship behaviors and mentorship quality associated with formal mentoring programs: Closing the gap between research and practice. *Journal of Applied Psychology*, 91(3), 567. <https://doi.org/10.1037/0021-9010.91.3.567>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191. <https://doi.org/10.1037/0033-295x.84.2.191>
- Berk, R. A., Berg, J., Mortimer, R., Walton-Moss, B., & Yeo, T. P. (2005). Measuring the effectiveness of faculty mentoring relationships. *Academic Medicine*, 80(1), 66–71. <https://doi.org/10.1097/00001888-200501000-00017>
- Centers for Disease Control and Prevention. (2022, November 16). *Program evaluation home – CDC*. Centers for Disease Control and Prevention. Retrieved January 23, 2023, from <https://www.cdc.gov/evaluation/>
- Crisp, G. (2009). Conceptualization and initial validation of the College Student Mentoring Scale (CSMS). *Journal of College Student Development*, 50(2), 177–194. <https://doi.org/10.1353/csd.0.0061>
- Crisp, G., & Cruz, I. (2009). Mentoring college students: A critical review of the literature between 1990 and 2007. *Research in Higher Education*, 50(6), 525–545. <https://doi.org/10.1007/s11162-009-9130-2>
- Dominguez, N. (2012). *Mentoring unfolded: The evolution of an emerging discipline* [Doctoral dissertation, University of New Mexico]. Digital Repository. https://digitalrepository.unm.edu/oils_etds/6
- Dominguez, N. & Kochan, F. (2020). Defining mentoring: An elusive search for meaning and a path for the future. In B. J. Irby, J. N. Boswell, L. J. Searby, F. Kochan, R. Garza & N. Abdelrahman (Eds.), *The Wiley international handbook of mentoring* (pp. 3–18). WILEY Publications, Inc.
- Dreher, G. F., & Ash, R. A. (1990). A comparative study of mentoring among men and women in managerial, professional, and technical positions. *Journal of Applied Psychology*, 75(5), 539–546. <https://doi.org/10.1037/0021-9010.75.5.539>
- Eby, L. T. (2019, October 25). *Creating a mentoring research project*. Post-conference workshop presented at the 12th Annual Mentoring Conference, Albuquerque, New Mexico.

Eby, L. T., Allen, T. D., Evans, S. C., Ng, T., & DuBois, D. L. (2008). Does mentoring matter? A multidisciplinary meta-analysis comparing mentored and non-mentored individuals. *Journal of Vocational Behavior*, 72(2), 254–267. <https://doi.org/10.1016/j.jvb.2007.04.005>

Ensher, E. A., & Murphy, S. E. (1997). Effects of race, gender, perceived similarity, and contact on mentor relationships. *Journal of Vocational Behavior*, 50(3), 460–481. <https://doi.org/10.1006/jvbe.1996.1547>

Fleming, M., House, S., Hanson, V. S., Yu, L., Garbutt, J., McGee, R., Kroenke, K., Abedin, Z., & Rubio, D. (2013). The mentoring competency assessment: Validation of a new instrument to evaluate skills of research mentors. *Academic Medicine*, 88(7), 1002–1008. <https://doi.org/10.1097/acm.0b013e318295e298>

Gershenfeld, S. (2014). A review of undergraduate mentoring programs. *Review of Educational Research*, 84(3), 365. <https://doi.org/10.3102/0034654313520512>

Gilbreath, B., Rose, G. L., & Dietrich, K. E. (2008). Assessing mentoring in organizations: An evaluation of commercial mentoring instruments. *Mentoring & Tutoring: Partnership in Learning*, 16(4), 379–393. <https://doi.org/10.1080/13611260802433767>

Grossman, J., & Mackenzie, F. J. (2005). The randomized controlled trial: Gold standard, or merely standard? *Perspectives in Biology and Medicine*, 48(4), 516–534. <https://doi.org/10.1353/pbm.2005.0092>
Hunt, D. M., & Michael, C. (1983). Mentorship: A career training and development tool. *Academy of Management Review*, 8(3), 475–485. <https://doi.org/10.5465/amr.1983.4284603>

Hurtado, S., Han, J. C., Sáenz, V. B., Espinosa, L. L., Cabrera, N. L., & Cerna, O. S. (2007). Predicting transition and adjustment to college: Biomedical and behavioral science aspirants' and minority students' first year of college. *Research in Higher Education*, 48(7), 841–887. <https://doi.org/10.1007/s11162-007-9051-x>

Jacobi, M. (1991). Mentoring and undergraduate academic success: A literature review. *Review of Educational Research*, 61(4), 505–532. <https://doi.org/10.3102/00346543061004505>

Johnson, W. B., Rose, G., & Schlosser, L. Z. (2010). Student-faculty mentoring: Theoretical and methodological issues. In T. D. Allen & L. T. Eby (Eds.), *The Blackwell handbook of mentoring: A multiple perspective approach*. John Wiley & Sons.

Kram, K. E. (1985). *Mentoring at work: Developmental relationships in organizational life*. Scott, Foresman and Company.

Martínez-Mesa, J., González-Chica, D. A., Bastos, J. L., Bonamigo, R. R., & Duquia, R. P. (2014). Sample size: How many participants do I need in my research? *Anais Brasileiros de Dermatologia*, 89(4), 609–615. <https://doi.org/10.1590/abd1806-4841.20143705>

McWilliams, A. (2017). Wake Forest University: Building a campus-wide mentoring culture. *Metropolitan Universities*, 28(3), 67–79. <https://doi.org/10.18060/21449>

Müller, T. (2008). Persistence of women in online degree-completion programs. *International Review of Research in Open and Distributed Learning*, 9(2), 1–18. <https://doi.org/10.19173/irrodl.v9i2.455>

National Mentoring Resource Center. (2016). *Resource assessment*. National Mentoring Resource Center. <https://nationalmentoringresourcecenter.org/resources/program-assessment/>

Noe, R. A. (1988). An investigation of the determinants of successful assigned mentoring relationships. *Personnel Psychology*, 41(3), 457–479. <https://doi.org/10.1111/j.1744-6570.1988.tb00638.x>

Nora, A., & Crisp, G. (2007). Mentoring students: Conceptualizing and validating the multi-dimensions of a support system. *Journal of College Student Retention: Research, Theory & Practice*, 9(3), 337–356. <https://doi.org/10.2190/cs.9.3.e>

Office of Human Subjects Research Protections. (n.d.). *Step 1: Do you need to submit to the IRB?* National Institutes of Health. Retrieved January 14, 2023, from [https://irbo.nih.gov/confluence/display/ohsrp/Step 1](https://irbo.nih.gov/confluence/display/ohsrp/Step+1)

O'Neil, J. M., & Wrightsman, L. S. (2001). The mentoring relationship in psychology training programs. In S. Walfish & A. K. Hess (Eds.), *Succeeding in graduate school: The career guide for psychology students* (pp. 113–129). Lawrence Erlbaum.

Ragins, B. R., & McFarlin, D. B. (1990). Perceptions of mentor roles in cross-gender mentoring relationships. *Journal of Vocational Behavior*, 37(3), 321–339. [https://doi.org/10.1016/0001-8791\(90\)90048-7](https://doi.org/10.1016/0001-8791(90)90048-7)

Ragins, B. R., & Scandura, T. A. (1999). Burden or blessing? Expected costs and benefits of being a mentor. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 20(4), 493–509. [https://doi.org/10.1002/\(sici\)1099-1379\(199907\)20:4<493::aid-job894>3.0.co;2-t](https://doi.org/10.1002/(sici)1099-1379(199907)20:4<493::aid-job894>3.0.co;2-t)

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54–67. <https://doi.org/10.1006/ceps.1999.1020>

Scandura, T. A. (1992). Mentorship and career mobility: An empirical investigation. *Journal of Organizational Behavior*, 13(2), 169–174. <https://doi.org/10.1002/job.4030130206>

Schäfer, M., Pander, T., Pinilla, S., Fischer, M. R., von der Borch, P., & Dimitriadis, K. (2015). The Munich-Evaluation-of-Mentoring-Questionnaire (MEMeQ)—a novel instrument for evaluating protégés' satisfaction with mentoring relationships in medical education. *BMC Medical Education*, 15(1), 1–8. <https://doi.org/10.1186/s12909-015-0469-0>

Simon, R. A., Aulls, M. W., Dedic, H., Hubbard, K., & Hall, N. C. (2015). Exploring student persistence in STEM programs: A motivational model. *Canadian Journal of Education*, 38(1), n1.

Tenenbaum, H. R., Crosby, F. J., & Gliner, M. D. (2001). Mentoring relationships in graduate school. *Journal of Vocational Behavior*, 59(3), 326–341. <https://doi.org/10.1006/jvbe.2001.1804>

Tinoco-Giraldo, H., Torrecilla Sánchez, E. M., & García-Peñalvo, F. J. (2020). E-mentoring in higher education: A structured literature review and implications for future research. *Sustainability*, 12(11), 4344. <http://dx.doi.org/10.3390/su12114344>

Tinto, V. (1987). *Leaving college: Rethinking the causes and cures of student attrition*. University of Chicago Press.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). University of Chicago Press.

University of Florida. (2022, October 28). *Institutional Review Board: Home*. Institutional Review Board. Retrieved January 23, 2023, from <https://irb.ufl.edu/>

Webber, S., & Prouse, C. (2018). The new gold standard: The rise of randomized control trials and experimental development. *Economic Geography*, 94(2), 166–187. <https://doi.org/10.1080/00130095.2017.1392235>

Yukawa, M., Gansky, S. A., O'Sullivan, P., Teherani, A., & Feldman, M. D. (2020). A new mentor evaluation tool: Evidence of validity. *PLOS ONE*, 15(6), e0234345. <https://doi.org/10.1371/journal.pone.0234345>

Notes

1. 45 CFR 46 102.d (<https://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html>)
2. <https://irb.ufl.edu>
3. The theory of change found in Appendix A of Case Study 3 in Chapter 16 includes proposed outcomes for the mentors as well as the mentees. For the sake of simplicity in illustrating methodological rigor in this section, we have omitted mentor outcomes from these if/then statements.