Landscape of Assessments of Mentoring Relationship Processes in Postsecondary STEMM Contexts: A Synthesis of Validity Evidence from Mentee, Mentor, Institutional/Programmatic Perspectives

Paul R. Hernandez, Ph.D.

West Virginia University

Commissioned paper prepared for the National Academies of Sciences, Engineering, and Medicine Committee on the Science of Effective Mentoring in Science, Technology, Engineering, Medicine, and Mathematics (STEMM)

Executive Summary:

- The present report is a synthesis of the validation evidence on assessments of mentorship relationship processes in postsecondary science, technology, engineering, mathematics, and medical (STEMM) contexts. This research synthesis focuses on measurement validity evidence based on the: (A) content of assessments; (B) internal structure of assessments; (C) relations between mentoring processes; and (D) reciprocal feedback between the perspectives of mentors and mentees.
- The majority of assessments and validation evidence concern mentoring from the mentee perspective. Summarizing across assessments from the mentee perspective, validation evidence is relatively robust for *some* relationship processes (i.e., instrumental support received, psychosocial support received, and relationship quality) based on content, internal structure and relations among processes. However, validation evidence for any particular assessment used in STEMM contexts is relatively limited (i.e., most assessments had only one study providing one-or-more forms of validation evidence). In addition, there is limited or no validation evidence from the mentee perspective for assessments of role modeling and negative experience in STEMM contexts. Furthermore, there has been no validation evidence gathered on reciprocal feedback between mentees and mentors.
- Only a small number of assessments have been developed or used in STEMM contexts from the mentor perspective. Similar to above, validation evidence for any particular assessment is limited (i.e., typically a single study reporting one-or-more forms of validation evidence in a STEMM context). Although more assessments have been developed from institutional/programmatic perspectives – the validation evidence for particular assessments remains similarly sparse (i.e., typically only one study reporting validation evidence).

Furthermore, assessments from institutional/programmatic perspectives typically only report validation evidence based on content. Therefore, validation evidence from mentor and institutional/program evaluation perspectives can be described as emerging, but limited in terms of evidence based on content, internal structure, relations with other variables, and reciprocal exchanges.

- Recommendations include the following:
    - Additional/future studies of mentoring in STEMM contexts should make more widespread use of established standards in educational and psychological measurement to develop robust validation evidence for individual assessments (AERA, APA, & NCME, 2014). Studies that use mentoring assessments should report one-or-more forms of validation evidence to help develop a more robust evidence base for a given assessment.
    - Additional studies of particular assessments are needed to develop more robust validation evidence based on content (particularly for role modeling and negative experience), internal structure, relations among mentoring processes, and reciprocal exchanges between mentor-mentee dyads.
    - Additional studies are needed to describe the degree to which developmental stage impacts the assessments of mentoring relationship processes. Mentoring relationships and assessments of those relationships may vary across developmental stages (i.e., undergraduate, graduate, postdoc, junior faculty, etc.). That is, developmental stage may influence: (A) which indicators are most salient/important; (B) the internal structure of an assessment; (C) the strength of associations among mentoring process; or (D) the characteristics of reciprocal feedback between mentor and mentee. Further

study in STEMM contexts (e.g., using the same assessment in multiple developmentally distinct samples, and/or following a single sample through multiple developmental stages) could illuminate the impact of development.

- Additional studies of existing assessments and/or development of new assessments that align content across perspectives are needed to develop our understanding of mentoring relationships and relational reciprocity (i.e., reciprocal exchanges) in STEMM contexts. Few assessments used in STEMM contexts align mentorship content across mentor and mentee perspectives (exceptions include the Mentor Competency Assessment, Negative Experiences survey [not currently validated in STEMM], and Working Alliance survey [mentor perspective not currently validated in STEMM]) (Eby, Butts, Lockwood, & Simon, 2004; Eby, Durley, Evans, & Ragins, 2008; Fleming et al., 2013; Schlosser & Gelso, 2001; Schlosser & Gelso, 2005).
- Additional studies using dyadic data analyses are needed to assess and describe the ways in which mentors and mentees experience interdependence and mutual influence. The concept of reciprocal exchanges between mentors and mentees is central to mentoring theory; yet, no validation evidence currently exists in STEMM contexts. This may be due, in part, to historical limitations in methods of assessing feedback loops. However, there are a number of promising relatively new measurement and statistical methods to assess reciprocal feedback loops in mentoring relationships via dyadic data analysis (Kenny, Kashy, & Cook, 2006).
- Additional studies describing and characterizing the development of mentor networks are needed to advance our understanding of mentoring relationships and to design policies to promote optimal mentoring relational outcomes in STEMM contexts.

Advancements in mentoring theory now recognize the potential importance of mentor support networks (Christou et al., 2017; de Janasz & Sullivan, 2004; Higgins & Kram, 2001; Montgomery, 2017). Recent advancements in statistical methods now allow for the assessment of mentor networks via social network analysis (Chariker, Zhang, Pani, & Rouchka, 2017; Scott, 2017).

1. Goals and scope of the present landscape review

The goal of this review was to provide a snapshot of the landscape of metrics, assessments, and methods used to research mentoring relationships in postsecondary educational science, technology, engineering, mathematics, and medical (STEMM) contexts. Although theoretical and operational definitions of mentoring vary across the literature (Jacobi, 1991), the present review defined mentoring as a developmental relationship between a more experienced person and a less experienced person (mentor and mentee, respectively), where the mentor provides support, guidance, and encouragement with the aim of enhancing the mentee's personal and/or professional development (Eby et al., 2013; Jacobi, 1991; Kram, 1985). Theoretical models of mentoring similarly vary across the literature; however, recent theoretical and empirical evidence supports a process-oriented model of mentoring, Figure 1 (Eby et al., 2013).

The process-oriented model depicted in Figure 1 shows that personal, contextual, and relational *inputs* shape the characteristics of mentoring relationship *processes*[1], and these relationship processes influence cognitive, emotional, and behavioral *outputs*. Outputs of mentoring relationships (i.e., the benefits of mentorship) are frequently of high interest in substantive studies of mentoring programs and reviews of the mentoring literature. Outputs from mentoring in STEMM contexts vary widely across the literature; however, some potential outputs include psychological processes (e.g., self-efficacy), learning or skill development (e.g., disciplinary knowledge, understanding the nature of science), scholarly achievement (e.g., GPA, presentations, publications), and enhanced career aspirations and advancement (e.g., persistence).

[1] Figure 1 has been slightly modified from the process-oriented model presented by Eby and colleagues (2013), in that the figure includes negative mentoring experiences and role modeling as distinct relationship processes (Eby and colleagues [2013] only included instrumental psychosocial support). The addition of negative mentoring experiences and role modeling as distinct relationship processes reflect current theory, measurement, and evidence on the nomological network of mentoring relationship processes (Eby et al., 2004; Eby et al., 2008; Hernandez et al., 2018)

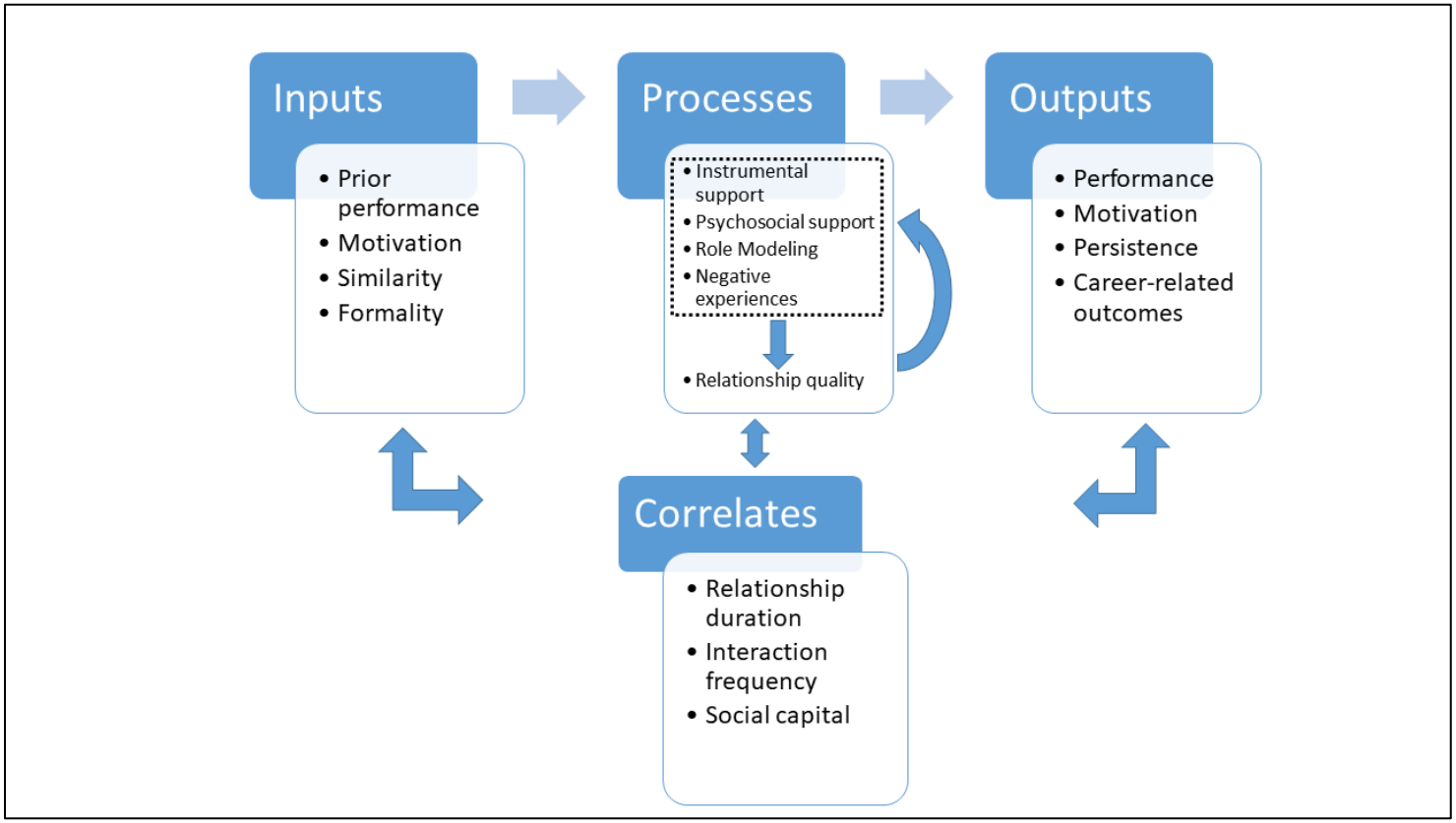

Figure 1. Process-oriented model of mentoring (adapted from Eby et al., 2013).

Readers interested in outputs (inside and outside of STEMM contexts) are encouraged to consult one of a number of thorough reviews on the topic (Crisp & Cruz, 2009; Eby et al., 2013; Gershenfeld, 2014; Ghosh, 2014; Jacobi, 1991; Pfund, Byars-Winston, Branchaw, Hurtado, & Eagan, 2016; Sadler, Burgin, McKinney, & Ponjuan, 2010; Syed, Azmitia, & Cooper, 2011).

Consistent with the goals for this study, the present review focuses on the metrics, assessments, and methods for studying mentoring relationship processes, rather than focusing on the inputs to or outputs from relationship processes. The following review will provide the following:

- working definitions of mentoring from the mentee, mentor, and institutional/programmatic perspective (section 2);
- an outline of the measurement validation framework, methods, and procedures used to conduct the present review (section 3);
- a summary of the strengths and weaknesses in the validation evidence-base for assessments used in STEMM contexts across different perspectives as well as, recommendations for further research on assessments of mentorship (section 4);
- and a discussion of the emerging approaches for measuring mentorship (section 5).

2. Defining mentoring relationship processes

Theoretical and empirical advances in the process-oriented model indicate that mentoring relationships can involve at least four distinct support functions (Figure 1, Processes): instrumental support, psychosocial support, negative experiences, and role modeling. Mentoring includes active functions, such as instrumental support (e.g., sponsorship, coaching, exposure and visibility, protection, and challenging work assignments), psychosocial support (e.g.,

acceptance, counseling, and friendship), and negative experiences (e.g., mismatch within the dyad, distancing behavior, manipulative behavior, lack of expertise, and general dysfunctionality) (Eby et al., 2004; Eby et al., 2008; Kram, 1985). Mentoring also involves passive functions, such as role modeling, wherein a mentor serves as an inspirational example of the norms, attitudes, and behaviors necessary to achieve success (Lockwood & Kunda, 1997). Benevolent mentoring support functions (i.e., instrumental support, psychosocial support, & role modeling) promote relationship quality (e.g., overall relationship satisfaction, liking, effectiveness) (Kram, 1985); whereas negative experiences diminish relationship quality. Relationship quality, in turn, reciprocally influences future levels of mentor support functions (Eby et al., 2013).

To date, much of the evidence on the nomological network of mentor relationship processes comes from studies outside of post-secondary educational STEMM contexts (Eby et al., 2013; Ghosh, 2014). A nomological network is a description of the theoretically expected pattern of associations among the constructs – in this case, theory suggests that benevolent mentorship processes (i.e., instrumental support, psychosocial support, & role modeling) should be positively correlated with one-another and relationship quality; whereas negative mentoring experiences should be negatively correlated the other mentorship processes. Within STEMM contexts, the few investigations of mentoring relationship processes have generally found patterns of association that are consistent with those in other contexts (Dilmore et al., 2010; Fleming et al., 2013; Haeger & Fresquez, 2016; Hernandez, Estrada, Woodcock, & Schultz, 2016; Hernandez et al., 2018; Prime, Bernstein, Wilkins, & Bekki, 2015; Rice et al., 2009; Tenenbaum, Crosby, & Gliner, 2001). However, given the small number of investigations within

STEMM contexts, it is reasonable to expect that context-based variations in the model have yet to be systematically documented.

2.1 Mentee, mentor, institutional, and program evaluation perspectives

An important caveat for the study of mentoring relationships concerns the distinguishability of mentor and mentee perceptions of their relationship. The extant literature has largely focused on the mentee's perspective of relationship processes (i.e., perceptions of mentoring received); however, mentor perceptions of mentoring support provided may (or may not) correspond or align with mentee perceptions. Theoretical models of close relationships suggest that reciprocal exchanges between relationship partners are nearly universal (Brown, 1991; Fiske, 1992). Recent theoretical work integrating theories of mentoring and self-regulated learning provides a framework for understanding and researching the reciprocal / cyclical feedback between mentor and mentee (Schunk & Mullen, 2013). Schunk and Mullen (2013) proposed a reciprocal mentoring model wherein mentor behaviors influence mentee self-regulatory perceptions and behavior, which, in turn influences mentor self-regulatory perceptions and future behaviors. Recent empirical evidence from the mentoring literature hints at reciprocal feedback loops between mentor and mentee perceptions (Griese, McMahon, & Kenyon, 2016). In their study of faculty-undergraduate summer research mentoring dyads, Griese and colleagues (2013) found that some aspects of mentor perceptions (e.g., importance of autonomy [mentorship input]) positively influenced mentee self-perceptions (e.g., research autonomy [mentorship output]), but mentee perceptions were not found to influence mentor perceptions. Although Griese and colleagues focused on mentorship inputs and outputs, rather than mentorship relationship processes, mentoring theory would suggest that mentor perceptions of support provided should influence mentee perceptions of support received and vice-versa.

A final element of the present review concerns the distinguishability of the institutional and program evaluation perspective of mentoring relationship processes. Institutional and programmatic perspectives are typically focused on the contextual inputs to mentor-mentee relationship formation and the associated relationship processes (e.g., institutional and programmatic efforts to increase the quantity and quality of mentoring relationships) (Lewis et al., 2016; Pfund, Pribbenow, Branchaw, Lauffer, & Handelsman, 2006; Silet, Asquith, & Fleming, 2010; Sorkness, Pfund, Asquith, & Drezner, 2013). Although institutional/programmatic perspectives frequently gather data from mentors or mentees, the stated purpose(s), framing, and use of associated assessments reflect an institutional perspective. That is, assessments designed for institutional purposes tend focus on changes in relationship formation (e.g., increases in numbers of students being mentored) or changes in mentoring skill development (e.g., increases in mentoring self-efficacy) before or after a program has been implemented, rather than on the quality of relationships among particular mentor-mentee dyads.

3. Synthesis of Validation Evidence Methods and Procedures

3.1 Search and inclusion criteria

This landscape review builds on a number of excellent multidisciplinary reviews of both mentoring and assessments of mentoring processes (Abedin et al., 2012; Allen, Eby, O'Brien, & Lentz, 2008; Chen, Watson, & Hilton, 2016; Crisp & Cruz, 2009; George & Neale, 2006; Gershenfeld, 2014; Huskins et al., 2011; Jacobi, 1991; Lee et al., 2012; Meagher, Taylor, Probsfield, & Fleming, 2011; Pfund et al., 2016). For example, Chen and colleagues (2016) reviewed assessments of mentoring relationships across educational, social/political sciences, biological/natural sciences, and nursing domains in education and professional contexts. The

present review incorporated information on assessments of mentoring relationships identified in the prior literature and conducted an updated search for mentoring assessments in the Web of Science database. In addition, cited reference searches were performed on all assessments identified to capture their use in STEMM contexts.

The present review extends the literature by focusing on evidence of measurement validity within postsecondary STEMM contexts to the summer of 2018. Therefore, the current review has been restricted to studies that included the following elements: (A) empirical (i.e., gathered data); (B) studied mentoring relationships in a postsecondary educational context (i.e., undergraduate student, graduate student, or postdoctoral mentees); (C) studied mentoring relationships in STEMM contexts (i.e., all or majority of mentees were from a STEMM discipline); (D) measured mentoring relationship processes from mentee, mentor, and/or institutional/program evaluation perspectives; (E) provided sufficient description of the sample, procedures, and measures to allow for replication (i.e., explicit details provided in the research article so that other researchers could replicate); and (F) provided one-or-more sources of measurement validation evidence.

3.2 Validation evidence framework

The present review highlights and synthesizes evidence of measurement validity for assessments of mentoring relationship processes in postsecondary STEMM contexts. The unified view of measurement validity defines validation as the theoretical and empirical process of evaluating the degree to which interpretations of, and actions based upon test scores are appropriate and reasonable (AERA et al., 2014; Messick, 1995). Under the unified view, all forms of validity evidence are elements of *construct validity* – that is, they provide evidence of the degree to which an assessment measures the construct it has been designed to measure

(Messick, 1995). Measurement validation is an extended process involving the synthesis of multiple sources of evidence, including the appropriateness and representativeness of the content in an assessment (*content*); the alignment between construct and process by which participants respond on an assessment (*response process*), the degree to which items on an assessment relate to constructs in a pattern consistent with theory and the proposed interpretations of scores from the assessment (*internal structure*), the degree to which scores from the assessment relate to other variables in a pattern consistent with theory (*relations to other variables*), and the appropriateness of the proposed interpretations of the scores from the assessment (*consequences*) (AERA et al., 2014).

The present review highlights validation evidence based on an assessment of content, internal structure, and relations among mentoring relationship processes, as well as evidence of internal consistency reliability for assessments used in STEMM contexts. A variety of methods can be used to provide validation evidence based on content. These include aligning item content with specified standards (typical of achievement tests), or identifying key facets of the construct from a literature review, expert judgement, systematic observation of behavior, focus groups, and/or interviews methods (AERA et al., 2014; McKenzie, Wood, Kotecki, Clark, & Brey, 1999; Rubio, Berg-Weger, Tebb, Lee, & Rauch, 2003; Vogt, King, & King, 2004). Evidence of internal structure is typically provided through the use of factor analytic (exploratory or confirmatory) or item response theory methods (AERA et al., 2014). Finally, evidence of relations with other variables can be provided through a variety of correlational methods (AERA et al., 2014).

3.3 Sample of assessments of mentoring relationship processes used in STEMM contexts

Thirty-five assessments of mentoring relationship processes in post-secondary educational STEMM contexts from mentee, mentor, or institutional/program evaluation perspectives were identified (Appendices 1-3). Across the literature in STEMM contexts, the majority of assessments have focused on measuring characteristics of the mentoring relationship from the mentee's perspective (number of assessments [*k*] = 22, 63%, Appendix 1), with relatively few assessments focusing the mentor's perspective of the relationship (*k* = 3, 8.5%, Appendix 2). Approximately one-quarter of the assessments have focuses on characteristics of mentoring relationships from an institutional or program evaluation perspective (*k* = 10, 28.5%, Appendix 3).

Assessments of a mentoring relationship from the mentee's perspective were adapted to or developed in a variety of post-secondary educational STEMM contexts (number of studies using assessments in a STEMM context [*n*] n_{Mentee} = 22, Appendix 1). Most studies of mentee perceptions of a mentoring relationship focused on undergraduate (59%) or graduate (36%) student mentees, with fewer focused on postdocs (5%). Almost half of the studies focused on the perspectives of mentees from historically underrepresented groups in STEMM disciplines (45%). Assessments from the mentor's perspective (n_{Mentor} = 3) focused on university faculty (66.5%) or graduate student and postdoc (33.5%) perceptions of mentoring relationship with undergraduates (Appendix 2). Finally, assessments of mentoring relationships from institutional/program evaluation perspectives ($n_{I/PE}$ = 8) have drawn on the perceptions of institutional staff members that run mentoring programs (38%) or faculty mentors involved in programs (63%, Appendix 3).

4. Landscape Review Results

The quantity and quality of validation evidence varies substantially both across perspectives (mentee, mentor, institutional & program evaluation) and within specific assessments from each perspective. Figure 2 summarizes the validation evidence based on assessment content, internal structure, and relationships among constructs within the process-oriented model of mentoring (Eby et al., 2013). That is, Figure 2 summarizes validity evidence for mentoring relationship processes (i.e., Figure 1, "Processes") in STEMM contexts. A synthesis of validation evidence based on content, internal structure, and relations with other mentorship processes follows.

4.1 Validation evidence based on assessment content

The process of gathering validation evidence begins by determining the degree to which the content of an assessment adequately captures the relevant facets of the construct it purports to measure. In the present case, the content of assessments aimed at measuring mentoring relationship processes (i.e., facets of instrumental support, psychosocial support, role modeling, negative experiences, and institutional/program support for mentoring) are summarized in Figure 2 within boxes titled "Indicators of…" on the right side of the figure. Individual indicators (e.g., "Access to resources") were drawn from assessments identified in Appendices 1-3.

A relative strength of the mentoring in STEMM literature concerns the quantity of validation evidence based on the content of assessments. Studies on assessments of mentoring processes have commonly used literature reviews, expert judgement, and/or the analysis of themes from interviews or focus groups to generate and validate assessment content.

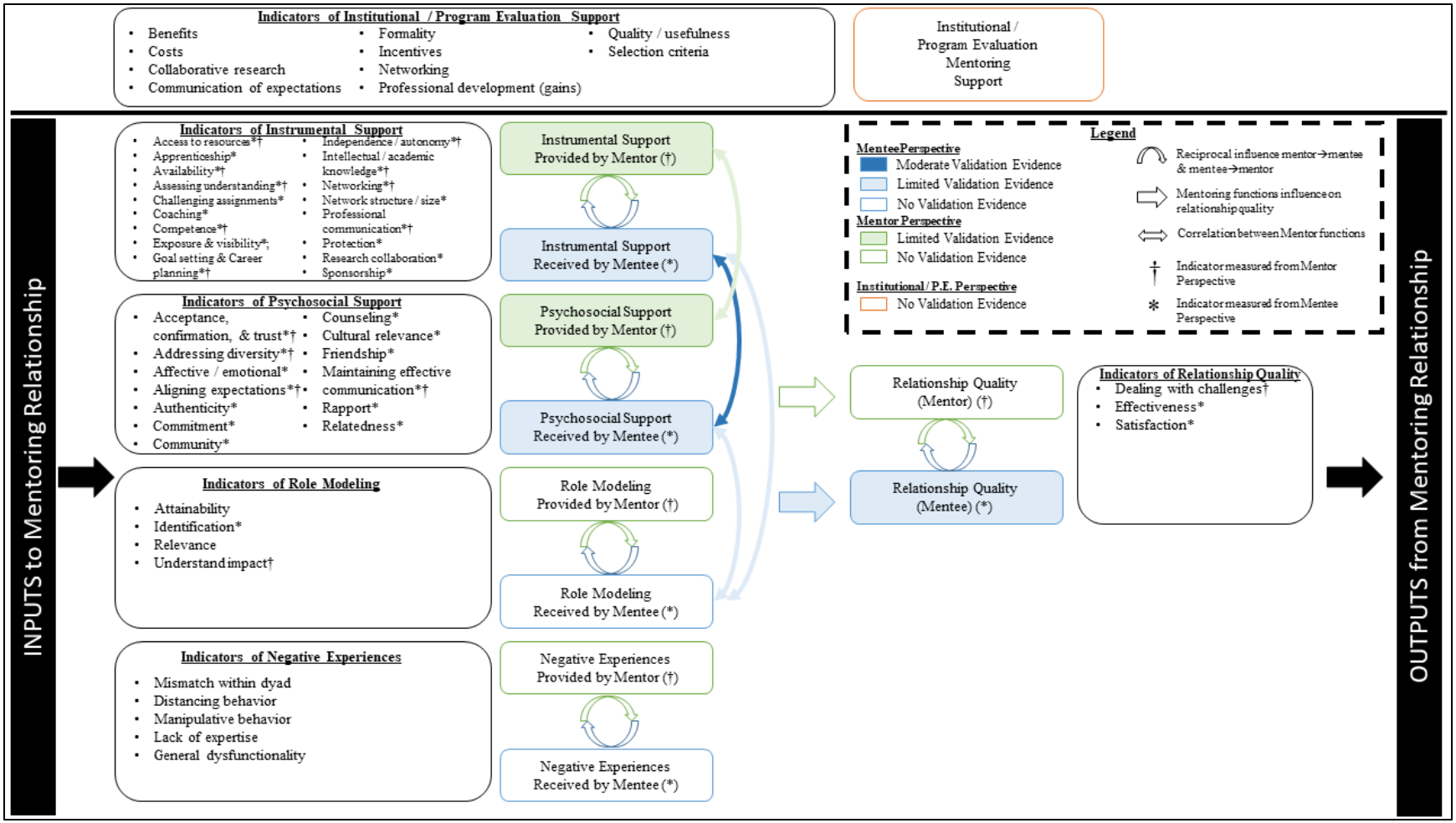

Figure 2. Synthesis of mentoring relationship processes validation evidence in postsecondary STEMM contexts.
Notes: The solid line separating institutional from mentor/mentee perspectives represents that there is little/no evidence connecting these perspectives. For the sake of simplicity, double-headed arrows were omitted where no evidence of a correlation has been reported in STEMM contexts.

The resulting assessments include diverse indicators of instrumental and psychosocial support from mentor and mentee perspectives (e.g., Figure 2, Indicators of Instrumental Support), and numerous indicators of support for mentoring from the institutional/programmatic perspective (Figure 2, Indicators of Institutional/Program Evaluation Support).

4.1.1 Mentee and mentor perspectives. Across the landscape, assessments from the mentee perspective (Figure 2, *) tap a wide variety of facets of instrumental support received (17 facets), psychosocial support received (14 facets), and relationship quality (2 facets). Facets range from general support functions that would manifest across contexts (e.g., goal setting) to support functions that are specific to STEMM contexts (e.g., research collaboration). Despite the fact that there are far fewer assessments from the mentor perspective (Figure 2, †) used in STEMM contexts, these assessments tap a relatively wide variety of facets of instrumental support (9 facets) and psychosocial support (4 facets). Facets of mentoring support provided range from general support functions to aspects specific to STEMM contexts (e.g., fostering research independence).

Two themes are evident from the landscape review. First, publication trends in the mentee literature indicate a shift away from using or adapting assessments developed in non-STEMM (typically corporate) contexts. Instead, recent publication trends indicate movement toward developing new assessments specifically for post-secondary and STEMM contexts. This shift toward contextualized assessments of mentoring has proven fruitful in that newly developed assessments appear to have identified facets of instrumental and psychosocial support that may be uniquely relevant to postsecondary STEMM (e.g., academic subject knowledge, aligning research expectations, research collaboration, Appendix 1). The cost of this approach to tailoring assessments to the STEMM context is that the field must now engage in the intensive effort

required to generate and evaluate validation evidence for the new instruments (AERA et al., 2014). By contrast, assessments from mentor and institutional/programmatic perspectives were developed in (rather than adapted to) postsecondary STEMM contexts. That is, this review found no evidence of assessments being adapted from non-STEMM contexts for use in postsecondary STEMM. As above, the potential benefits and costs of tailoring assessments to postsecondary STEMM contexts apply.

Two examples may illustrate the benefits and costs of adaptation to, or development in postsecondary STEMM contexts. The adaptation of the Global Measure of Mentoring Practices (GMMP) (Dreher & Ash, 1990) is an excellent example of adapting a measure designed for a non-STEMM context (Appendix 1). Dreher and Ash (1990) developed the GMMP as a global assessment of mentorship support received (i.e., mentee perspective) based on Kram's framework of nine mentor roles (Dreher & Ash, 1990; Kram, 1985). Based on their review of the literature and expert judgment, Tenenbaum and colleagues (2001) adapted the GMMP for use in postsecondary STEMM contexts by omitting two questions that were irrelevant to graduate students and adding four additional questions that related to disseminating research and exploring career options (Tenenbaum et al., 2001). The resulting adapted GMMP measures 10 facets of instrumental and psychosocial support (i.e., Kram's 9 mentoring roles plus research dissemination; a subset of all identified facets) that are applicable to mentee experiences postsecondary STEMM. The adaptation of the GMMP was efficient and relatively low in cost (i.e., made use of instrument with existing content validation evidence), but the benefits may be limited by lack of specificity to and omission of relevant facets unique to STEMM. By contrast, the development of the Mentoring Competency Assessment (MCA) (Fleming et al., 2013) is an excellent example of the rigorous and costly processes for gathering content validation evidence

for an assessment tailored to mentors and mentees in postsecondary STEMM research contexts (Appendix 1). The content validation process involved: (A) an extensive review of the mentoring assessments; (B) cognitive interviews with mentors and mentees in postsecondary STEMM research contexts; and (C) aligning assessment content to a framework and learning objectives for a mentor training program (Fleming et al., 2013; Handelsman, Pfund, Lauffer, & Pribbenow, 2005; Pfund et al., 2013; Pfund et al., 2006). The resulting MCA measures six facets of instrumental and psychosocial support (i.e., a subset of all identified facets) that are well specified for postsecondary STEMM research contexts. The development of the MCA was relatively costly, but the benefits may be signficant in that the MCA has a high degree of specificity to, and includes relevant facets unique to STEMM. Clearly, the decision to adapt or develop an assessment (and in particular, an assessment's content) for postsecondary STEMM is not trivial. The decision to adapt or develop content should be informed by a variety of factors, including the theory of mentoring (i.e., theory-guided selection of facets and nomological network), the required specificity, and the consequences of decisions that will be based on scores from assessments.

The second theme concerned theoretical consensus. As is true of the broader mentorship literature (Crisp & Cruz, 2009; Jacobi, 1991), there is little consensus on the most relevant or essential facets of instrumental or psychosocial support provided or received (i.e., mentor and mentee perspectives, respectively). This consensus issue stems from the lack of a common definition for mentoring, as well as from diverse theoretical perspectives on the roles that mentor and mentees play. The lack of consensus is evident in the content of assessments in that (A) any given assessment only taps a subset of the wider variety of facets of instrumental and

psychosocial support and (B) the lack of parity between facets of support assessed from mentee and mentor perspectives (Appendices 1-2).

4.1.2 Institutional/Programmatic perspectives. As above, assessments from the institutional perspective tap a variety of facets of structures ranging from perceived costs/benefits to professional development opportunities. However, to date, there is a paucity of theoretical or empirical work linking the content or facets of institutional support structures for mentoring to dyadic mentoring processes (e.g., perceptions of mentoring provided by a mentor to a mentee). Therefore, the current assessments of mentoring from institutional and programmatic perspectives do not align well with theoretical models of mentoring relationship processes (i.e., instrumental support, psychosocial support, role modeling, and negative experiences).

4.1.3 Gaps. The most notable gap concerns the scarcity of validation evidence based on content for relationship quality (mentor's perspective) and for role modeling or negative experiences (any perspective). The absence of robust content validation evidence for role modeling and negative experiences may be relatively easy to correct. That is, there is robust theoretical and empirical guidance on the relevant facets of role modeling (i.e., attainability of a role model's achievements, relevance of a role model, and identification with a role model) outside of STEMM contexts (Hoyt, Burnette, & Innella, 2012; Lockwood, 2006; Lockwood & Kunda, 1997). Similarly, there is robust theoretical and empirical content validation evidence on assessments of negative mentoring experiences outside of STEMM contexts (Eby et al., 2004; Eby et al., 2008). Furthermore, there is content validation evidence on the relevant facets of relationship quality in STEMM contexts from the mentee perspective (Byars-Winston, Branchaw, Pfund, Leverett, & Newton, 2015; Dennehy & Dasgupta, 2017; Ensher & Murphy, 1997; Hernandez et al., 2016). In the short-term, the process of gathering validation evidence

based on assessment content for role modeling, negative experiences, and relationship quality may be relatively easy to correct through the process of adaptation of existing instruments (e.g., Negative mentoring experiences scales for mentor and mentee by Eby and colleagues [2004, 2008]). In the long-term, it will likely be beneficial to conduct the types of phenomenological studies (e.g., interviews) required to fully elaborate the facets of these relationship processes as they manifest in postsecondary STEMM contexts.

4.2 Validation evidence based on internal structure

The strength of validation evidence based on internal structure for assessments of mentoring in STEMM varies widely across perspectives and constructs (in Figure 2, boxes coded with darker colors indicate evidence is more robust). The methods commonly used to validate an assessment's internal structure involved exploratory factor analysis, confirmatory factor analysis, and/or item response theory models.

4.2.1 Mentee and mentor perspectives. Consistent with the quantity and quality of validation evidence based on content, most of the validation evidence based on internal structure concerned instrumental and psychosocial support and relationship quality (i.e., mentor [Figure 2, light green boxes] and mentee perspectives [Figure 2, light blue boxes]). A moderate proportion of the assessments from the mentee perspective (12 of 22) and most of the assessments from the mentor's perspective (2 of 3) provided one form of evidence concerning internal structure (e.g., an exploratory factor analysis that produced the theoretically derived factor structure). Unfortunately, relatively few assessments have had more than one study yielding evidence on internal structure in postsecondary STEMM contexts, and where multiple sources of evidence do exist, internal structure is sometimes unstable (Appendices 1-3). For example, three studies with graduate student mentees in STEMM contexts tested the factor structure of the mentoring

functions and relationship satisfaction scales on the Survey of Doctoral Education (Golde & Dore, 2001; Noy & Ray, 2012; Rice et al., 2009) – each coming to a dramatically different conclusion about the internal structure. In another example, researcher's studying the internal structure of the GMMP identified a 3-factor structure in a sample of graduate students (Tenenbaum et al., 2001), but a 2-factor structure in a sample of undergraduate students (Hernandez et al., 2016). Therefore, the validation evidence base concerning internal structure within STEMM contexts is still emerging.

4.2.2 Institutional/Programmatic perspectives. No evidence of internal structure has been gathered in STEMM contexts from the institutional/programmatic perspective. Evidence of internal structure (e.g., factor analysis) requires a multi-indicator approach; however, assessments of institutional/programmatic supports for mentoring have relied on single-indicator approaches to measure constructs (i.e., each question measures a different construct, and indicators are not aggregated).

4.2.3 Gaps. Several gaps are evident from the landscape in STEMM contexts. First, the internal structure evidence base for any given assessment is relatively thin (typically only one study, when evidence was provided), and when multiple studies of internal structure have been reported there is evidence of variability of the internal structure. It is possible that some of the variability concerning internal structure may be due to developmental changes in the way support is provided or received; however, more studies across different developmental stages will be required to investigate this possibility. Second, to date, no studies have gathered evidence on the internal structure of assessment of role modeling or negative experiences (any perspective), or relationship quality (mentor perspective). Third, the validation evidence base from an institutional or program evaluation is highly limited due to the exclusive use of single-

indicator approaches to measure facets of support for mentoring, which eliminates the possibility of establishing an internal structure.

4.3 Validation evidence based on relations with other variables

As discussed above, the strength of validation evidence based on relations with other mentoring relationship processes varies widely across perspectives and constructs (in Figure 2, darker color-coded single- and double-headed arrows indicate evidence is more robust). The methods commonly used to validate relations among mentor relationship processes involved correlational, regression, and structural equation analyses.

4.3.1 Mentee and mentor perspectives. Most of the validation evidence based on relations with other variables concerned associations between instrumental support and psychosocial support (Figure 2, dark blue double-headed arrow) and, to a lesser extent, included relations with role models (Figure 2, light blue double-headed arrow), or relationship quality (Figure 2, light blue single-headed arrow). A moderate proportion of the assessments from the mentee perspective (14 of 22) and most of the assessments from the mentor's perspective (2 of 3) provided evidence for the positive associations among these relationship processes.

4.3.3 Gaps. Several gaps are apparent from the landscape study. First, there is a lack of studies examining relations that include role modeling or negative experiences (any perspective) or relationship satisfaction (mentor's perspective) in postsecondary STEMM contexts. Second, no validation evidence currently exists for the reciprocal relations (i.e., feedback loop) between mentee and mentor perspectives of support provided and received (e.g., in Figure 2, empty blue and green feedback loops connecting Instrumental Support Received by Mentee [*] and Instrumental Support Provided by Mentor [†]). Third, studies from institutional and

programmatic perspectives have been descriptive and have not evaluated associations among mentoring relationship support structures.

5. Discussion of New/Emerging Approaches

One of the major gaps in the mentoring literature inside and outside of STEMM contexts concerns assessment of the dynamic, reciprocal, and interdependent nature of mentor and mentee perspectives on their relationship. Although relationship theory has for some time pointed to the essential nature of reciprocal exchanges between relational partners (Brown, 1991; Fiske, 1992), mentoring theory has only recently begun to recognize the potential importance of reciprocal feedback (Schunk & Mullen, 2013). Thus, the field knows little about reciprocal exchanges between mentors and mentees in postsecondary STEMM contexts. However, relatively recent advances in statistical methodology now allow for characterizing reciprocal relationships through dyadic data analysis (Kenny, 1994; Kenny et al., 2006). Dyadic data analysis refers to a family of methods for quantifying the degree-to-which and ways-in-which relationship partners (e.g., mentor-mentee, romantic partners, or parent-child) are interdependent (i.e., the relationship is the unit of analysis). That is, dyadic data analysis could yield insights on how the thoughts, feelings, or behaviors of one relationship partner (e.g., the mentor) may to influence the thoughts, feelings, or behaviors of the other relationship partner (e.g., the mentee) and/or vise-versa. Improvements in the assessment and theory of reciprocal exchanges in mentoring may inform best practices in forming or maintaining mutually satisfying mentoring relationships that ultimately lead to beneficial outcomes. One study has used a dyadic approach to characterize reciprocal feedback between mentors and mentees in a STEMM research experience context

(Griese et al., 2016); however, wider use of dyadic methodology could inform provision and reception of mentoring support.

A second major gap in the assessment literature concerns the assessment of mentor networks. Theoretical advancements in mentoring theory point to the importance of diversified and developmental networks of mentoring relationships – particularly for individuals from historically underrepresented groups (Downing, Crosby, & Blake-Beard, 2005; Glessmer, Wang, & Kontak, 2012; Higgins, 2000; Higgins & Kram, 2001; Higgins & Thomas, 2001; Packard, 2003; Packard , Walsh, & Seidenberg, 2004). For example, several researchers have begun to measure and explore differential benefits of triadic mentorship structures (Aikens et al., 2017; Aikens et al., 2016; Morales, Grineski, & Collins, 2018). For example, Aikens and colleagues found that the "closed" triadic structure was uniquely beneficial for undergraduate research mentees, that is, where undergraduate mentees were connected to a faculty mentor (i.e., senior mentor) and a postgraduate mentor (i.e., a step-ahead mentor and also a mentee to the faculty). Beyond network structures, Christou and colleagues (2017) have suggested that network strength, diversity, interconnectivity, and connections to mentor(s) with power are important facets of mentor networks (Christou et al., 2017). Recent advancements in statistical methodology now allow researchers to measure and test hypotheses concerning mentor networks via social network analysis (Scott, 2017). Social network analysis refers to a family of statistical methods to characterize and quantify the patterns of relational connections among people in a community (Scott, 2017). Social network analysis could be used to simultaneously identify optimally beneficial mentor network structures (e.g., "closed" triads) and quantify individual differences in network strength, diversity, interconnectivity, and connections to power. Improvements in the assessment and longitudinal development of mentor networks could

illuminate mentor theory, inform best practices in the design and implementation of mentorship programs, and ultimately lead to better outcomes for mentees.

In addition to the two major gaps in the literature noted above, there are a number of smaller, but still significant gaps both across the mentoring in STEMM literature and for the literature on specific assessments. For example, there were variable and inconsistent approaches to reporting measurement validation evidence across the literature. Additional studies of mentoring in STEMM contexts should make more widespread use of established standards in educational and psychological measurement to develop robust validation evidence for individual assessments (AERA et al., 2014). Studies that use mentoring assessments should report one-or-more forms of validation evidence to help develop a more robust evidence base for a given assessment. Furthermore, additional studies of particular assessments are needed to develop more robust validation evidence based on content (particularly for role modeling and negative experience), internal structure, relations among mentoring processes, and reciprocal exchanges between mentor-mentee dyads. Theory and some evidence point to potential changes in the functioning of assessments across different stages of development (i.e., undergraduate, graduate, postdoc, junior faculty, etc.). That is, developmental stage may influence: (A) which indicators are most salient/important; (B) the internal structure of an assessment; (C) the strength of associations among mentoring process; or (D) the characteristics of reciprocal feedback between mentor and mentee. Further study in STEMM contexts (e.g., using the same assessment in multiple developmentally distinct samples, and/or following a single sample through multiple developmental stages) could illuminate the impact of development. Additional studies are needed to describe the degree to which developmental stage impacts the assessments of mentoring relationship processes. Finally, very few assessments used in STEMM contexts align mentorship

content across mentor and mentee perspectives (exceptions include the Mentor Competency Assessment, Negative Experiences survey [not currently validated in STEMM], and Working Alliance survey [mentor perspective not currently validated in STEMM]) (Eby et al., 2004; Eby et al., 2008; Fleming et al., 2013; Schlosser & Gelso, 2001; Schlosser & Gelso, 2005). This lack of alignment precludes understanding the dynamic and reciprocal ways in which mentor and mentees may experience interdependence and mutual influence. Additional studies of existing assessments and/or development of new assessments that align content across perspectives are needed to develop our understanding of mentoring relationships and relational reciprocity (i.e., reciprocal exchanges) in STEMM contexts.

Appendix 1.

Summary of Scales of Mentoring Used in STEMM from Mentee Perspective

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Mentoring Functions Scale [MFS, 29] (Noe, 1988)	Career support received, Psychosocial support received.	**Green & Bauer (1995); Paglis et al., (2006)** 233 first-year doctoral students in "Hard" sciences at a single university (24 departments) (Green & Bauer, 1995; Paglis, Green, & Bauer, 2006).	**Green & Bauer (1995); Paglis et al., (2006)** ***Content***: Original scale develop items to assess career and psychosocial mentor functions described by Kram (1985). Item wording adapted to STEMM context and one item (future advancement) omitted for contextual irrelevance. Complete list of relevant mentoring survey items provided. ***Internal Structure***: Exploratory factors analysis indicated two factors using Kaiser rule and scree plot; however, 8 of 20 items (40%) were cross-loaded. Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: N/A.
Mentor Role Instrument [MRI, 33] (Ragins & McFarlin, 1990)	Career roles (sponsor, coach, protector, challenger, and promoter); Psychosocial roles (friend, social associate, parent, role model, counselor, and acceptor).	**Dilmore et al., (2010)** 141 clinical and translational science trainees at an academic medical center (Dilmore et al., 2010).	**Dilmore et al., (2010)** ***Content***: Original scale develop items to assess career and psychosocial mentor functions described by Kram (1985). No wording changes from the original. Relationship quality and relationship effectiveness captured with a single item each. Complete list of relevant mentoring survey items provided. ***Internal Structure***: Confirmatory factor analysis used to test subscales within Career and Psychosocial roles (i.e., second-order model); evidence of adequate data-model fit. Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: Moderate-to-strong correlations between career support, psychosocial support, relationship quality, and relationship effectiveness.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Global Measure of Mentoring Practices [GMMP, 18] (Dreher & Ash, 1990)	Global measure of mentoring support functions received.	**<u>Tenenbaum et al. (2001)</u>** 189 graduate students (76% in STEMM disciplines) at a single university. (Tenenbaum et al., 2001) **<u>Hernandez et al., (2016)</u>** 253 African American undergraduates in STEM majors. (Hernandez et al., 2016) **<u>Hernandez et al. (In press)</u>** 203 undergraduates in STEM majors engaged in a summer research experience. (Hernandez et al., 2018)	**<u>Tenenbaum et al. (2001)</u>** ***Content***: Original GMMP sampled items from existing scales (e.g., Noe, 1988) to represent Kram's 9-mentor functions in a global fashion (i.e., single factor). Omitted 2 items deemed irrelevant to graduate students. Added 4 additional items related (3 related to disseminating research, 1 related to exploring career options) – 22 items. Complete list of relevant mentoring survey items provided. ***Internal Structure***: Used principle components analysis with Kaiser rule to determine three factors consisting of Psychosocial, Instrumental, and Networking functions. Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: Moderate correlations between instrumental support, psychosocial support, networking, and relationship quality. **<u>Hernandez et al., (2016)</u>** ***Content***: Selected 15 items deemed most relevant for undergraduates from GMMP and with cross-validation evidence from Tenenbaum et al. (2001). ***Internal Structure***: Used exploratory factor analysis, parallel analysis, and Velicer's minimum average partial test to determine two factors consisting of Psychosocial and Instrumental functions. Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: Moderate correlations with relationship quality. Unrelated to mentor-mentee research collaboration.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
			Hernandez et al. (In press) ***Content***: Selected 10 items deemed most relevant for undergraduates from GMMP and with cross-validation evidence from Tenenbaum et al. (2001). ***Internal Structure***: Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: Moderate correlations with relationship quality and role modeling.
Research Collaboration [10-5] (Green, 1991; Green & Bauer, 1995; Paglis et al., 2006)	Research Collaboration.	**Green & Bauer (1995); Paglis et al., (2006)** 233 first-year doctoral students in "Hard" sciences at a single university (24 departments). **Hernandez et al., (2016)** 253 African American undergraduates in STEM majors. (Hernandez et al., 2016)	**Green & Bauer (1995); Paglis et al., (2006)** ***Content***: Developed list of questions based on research on types of ways that faculty-graduate students collaborate on research projects (Green, 1991). Complete list of relevant mentoring survey items provided. ***Internal Structure***: N/R. ***Relationship with other mentoring relationship variables***: Small positive relationship with instrumental support, but uncorrelated with psychosocial support. **Hernandez et al., (2016)** ***Content***: N/A – used items as listed in Green & Bauer (1995). ***Internal Structure***: N/R. ***Relationship with other mentoring relationship variables***: Uncorrelated with psychosocial support, instrumental support, and satisfaction.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Mentor Satisfaction scale [3] (Ensher & Murphy, 1997)	Satisfaction.	**Hernandez et al., (2016)** 253 African American undergraduates in STEM majors. (Hernandez et al., 2016) **Hernandez et al. (In press)** 203 undergraduates in STEM majors engaged in a summer research experience. (Hernandez et al., 2018)	**Hernandez et al., (2016)** ***Content***: Minor adaptation of the instructions from the original scale, developed in non-academic mentoring context. ***Internal Structure***: Internal consistency reliability reported at >0.70. ***Relationship with other variables***: Moderate positive relationship with mentor support functions (instrumental & psychosocial). **Hernandez et al., (In press)** ***Content***: Minor adaptation of the instructions from the original scale, developed in non-academic mentoring context. ***Internal Structure***: Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: Moderate positive relationship with mentor support functions (instrumental & psychosocial).
Need Satisfaction Scale [9] (La Guardia, Ryan, Couchman, & Deci, 2000)	Autonomy, Competence & Relatedness.	**Lewis et al., (2016)** 137 underrepresented minority graduate / postdoc / junior faculty trainees at medical centers (Lewis et al., 2016).	**Lewis et al., (2016)** ***Content***: Adapted original scale to mentoring relationship. ***Internal Structure***: Used principle components analysis to determine three factors consisting of autonomy, competence, & relatedness need satisfaction received from traineeship mentor. Internal consistency reliability reported at >0.70.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Survey on Doctoral Education – Mentoring subscale [23] (Golde & Dore, 2001; Noy & Ray, 2012)	Mentoring functions, Relationship satisfaction.	**<u>Noy & Ray (2012)</u>** Secondary data analysis of Golde & Dore (2001) study of 4,114 doctoral students with identified advisors/mentors at 27 universities (64% in STEMM disciplines) (Noy & Ray, 2012). **<u>Rice et al., (2009)</u>** 367 international doctoral students (56% in STEMM disciplines) from a single university (Rice et al., 2009). **<u>Curtin et al., (2016)</u>** 848 graduate students in 26 departments (63% in STEMM disciplines) at a single university (Curtin, Malley, & Stewart, 2016).	**<u>Noy & Ray (2012)</u>** ***Content***: Analyzed only items related to support functions – did not analyze items related to satisfaction. ***Internal Structure***: Used principle components analysis to determine 6-factors consisting of *Affective*, *Instrumental*, *Intellectual*, *Exploitive*, *Available*, and *Respectful*. Methods of determining factor structure not mentioned and evidence of over-extraction / factor splitting. Internal consistency reliability reported at >0.70 for *Affective*, *Instrumental*, *Intellectual*, & *Available*. ***Relationship with other mentoring relationship variables***: N/R. **<u>Rice et al., (2009)</u>** ***Content***: Analyzed a mixture of 8 items related to satisfaction (4) and support function (4). Decision on which items to draw from the satisfaction and support function scales based on author judgement of item content related to satisfaction. ***Internal Structure***: Used exploratory factor analysis to determine single-underlying satisfaction scale (after dropping two items). Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: High correlations between the three subscale scores.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
			<u>Curtin et al., (2016)</u> ***Content***: N/A – scale unchanged from original. ***Internal Structure***: Used exploratory factor analysis retained 19 of 23 items and extracted three factors labeled instrumental, psychosocial, and sponsorship support (note that the evidence for sponsorship support was poor – doubled loaded items allowed to form this third factor). No information on factor extraction criteria. Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: N/R.
Working Alliance in Advisor–Advisee Relationships [AWAI, 29] (Schlosser & Gelso, 2001)	Rapport; Apprenticeship; Identification-Individuation.	**<u>Rice et al., (2009)</u>** 367 international doctoral students (56% in STEMM disciplines) from a single university (Rice et al., 2009). **<u>Prime et al., (2009)</u>** 293 female doctoral students in STEMM disciplines. (Prime et al., 2015) **<u>Aikens et al. (2016, 2017)</u>** 842 undergraduates from 50 institutions across the U.S. All participants reported having completed one-or-more semesters (or summers) of research and were currently working with both a faculty and postgraduate on a research project (Aikens et al., 2017; Aikens et al., 2016).	**<u>Rice et al., (2009)</u>** ***Content***: N/A – scale unchanged from original. ***Internal Structure***: Used confirmatory factor analysis to test the three factor structure. Results provided evidence of acceptable data-model fit for three factor structure. All internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: Moderate to high correlations among the support functions and between support functions and satisfaction.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
			Prime et al., (2009) ***Content***: Added 17-research generated items to the AWAI based on author's interest in specific mentoring types of mentoring support that may be specific to women in STEMM. ***Internal Structure***: Used exploratory factor analysis with screen plot and "interpretability" of factors to guide factor extraction. Extracted 3-factors related, but chose only to interpret first 2-factors: psychosocial and instrumental support, as the 3rd factor appeared to capture reverse-scored items. Resulting 2-factor structure contains only 12 of the 46 items. All internal consistency reliability reported at >.70 (not reported for 3rd factor). ***Relationship with other mentoring relationship variables***: Moderate to high correlations among the support functions and between support functions and satisfaction. **Aikens et al. (2017)** ***Content***: N/A – unchanged from original. ***Internal Structure***: Used confirmatory and exploratory factor analysis on the Rapport subscale. Determined that the three negatively worded items on the subscale formed their own factor. Internal consistency reliability reported at >.70. ***Relationship with other mentoring relationship variables***: Moderate positive correlations between *closed triad type*, frequency of interaction with mentors, and rapport (i.e., psychosocial support).

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Mentorship Effectiveness Scale [12] (Berk, Berg, Mortimer, Walton-Moss, & Yeo, 2005)		**Berk et al., (2005)** N.A. – no data collected. **Rorrer (2016)** 226 undergraduates enrolled in NSF funded REU programs for summer research (Rorrer, 2016).	**Berk et al., (2005)** ***Content***: Developed set of items based on review of the literature and committee review of the content of the items. ***Internal Structure***: N/A – no data collected. ***Relationship with other mentoring relationship variables***: N/A – no data collected. **Rorrer (2016)** ***Content***: N/A – content unchanged from original. ***Internal Structure***: Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: N/A.
College Student Mentoring Scale [CSMS, 25] (Crisp, 2009; Crisp & Cruz, 2010)	Psychological and emotional support, Goal setting and career paths, Academic subject knowledge support, & Role model.	**Peltz & Raymond (2016)** 249 undergraduate students in nursing major (Peltz & Raymond, 2016).	**Peltz & Raymond (2016)** ***Content***. N/A – no changes made to survey content. ***Internal Structure***: Internal consistency reliability reported for benefit of mentoring at >0.70. ***Relationship with other mentoring relationship variables***: N/R.
Student-Faculty Interaction [6, 9] (Chang, Sharkness, Hurtado, & Newman, 2014; Eagan et al., 2013; Sharkness,	Global measure of support and satisfaction.	**Sharkness et al. (2010)** Undergraduate students participating in the Higher Education Research Institute "Your First College Year" survey (YFCY) or "College Senior Survey" (CSF). Sample size not reported.	**Sharkness et al. (2010), Eagan et al. (2013), Chang et al. (2014)** ***Content***. Developed item pool based on Astin's involvement theory (1984, 1999) and Weidman's model of college student socialization (1989). First year students asked 8 questions about quality, frequency, and level of satisfaction with their interaction with faculty (YFCY). College seniors asked 9 questions about the frequency with which faculty provided mentorship, support, and guidance (CSF). Complete list of relevant mentoring survey items provided.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
DeAngelo, & Pryor, 2010)		**Eagan et al. (2013)** Longitudinal subsample 4,152 students from 219 institutions that completed the YFCY (2004) and CSF (2008) surveys. Subsample selected because they had indicated an interest in a STEM-related degree in 2004. **Chang et al. (2014)** Longitudinal subsample 3,670 students from 217 institutions that completed the YFCY (2004) and CSF (2008) surveys. Subsample selected because they had indicated an interest in a STEM-related degree in 2004.	***Internal Structure***: Only reported internal structure for YFCY measure of student-faculty interaction. Used iterative exploratory factor analyses to determine a one-factor solution. 2-items were removed from the analysis due to poor item-characteristics. Used item response theory (IRT) based graded response model on the remaining 6-items to characterize item parameters and estimate participant scores. ***Relationship with other mentoring relationship variables***: N/R.
Role Model Identification [4] (Hoyt et al., 2012)	Role model identification	**Hernandez et al. (In press)** 203 undergraduates in STEM majors engaged in a summer research experience. (Hernandez et al., 2018).	**Hernandez et al. (In press)** ***Content***: Four indicators align with principles of relevance and attainability described in the role modeling literature. Adapted instructions to focus on primary faculty mentor. ***Internal Structure***: Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: Moderate to high correlations between role model identification, support functions, and satisfaction.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Mentoring Competency Assessment [MCA, 26] (Fleming et al., 2013; Pfund et al., 2013; Pfund et al., 2014)	Maintaining effective communication, Aligning expectations, Assessing understanding, Addressing diversity, Fostering independence, & Promoting professional development.	**Fleming et al. (2013) & Pfund et al. (2014)** 283 mentor (faculty) –mentee (undergraduate through faculty) pairs from 16 universities involved in mentoring in STEMM training program.	**Fleming et al. (2013)** ***Content***: Aligned 26-indicators with 6-part framework and learning objectives of Mentor Training for Clinical and Translational Researchers workshop (Pfund et al., 2013). Pilot testing used cognitive interviews with mentors and mentees to assess the instrument's consistency. Provide complete list of relevant mentoring survey items. ***Internal Structure***: Confirmatory factor analysis conducted on the hypothesized 6-factor structure. Data-model fit indices showed less-than good-fit. Internal consistency reliability reported at >0.70 for Aligning expectations, Assessing understanding, Addressing diversity, & Fostering independence. Less than 0.70 for Maintaining effective communication & Promoting professional development. ***Relationship with other variables***: Moderate to high correlations among the six factors.
Mentor effectiveness scale [26] (Byars-Winston et al., 2015)	Effectiveness	**Byars-Winston et al. (2016)** 214 undergraduate researchers involved in a summer research opportunity program. Sample were 65% female and 77% were from underrepresented racial / minority groups.	**Byars-Winston et al. (2016)** ***Content***: Survey items based on pre-existing instruments used with undergraduate researchers. Provide sample of relevant mentoring survey items. ***Internal Structure***: Exploratory factor analysis conducted on 26 items. One item dropped based on low association with factor. Compared data-model fit of 1, 2, and 3 factor solutions. Authors concluded that one-factor provided adequate and parsimonious fit (model fit comparisons not provided). Reported data-model fit indices for single-factor solution showed good fit. Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: N/A.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Mentoring Structure, Motivation, & Effectiveness [32] (McGinn et al., 2015)	Mentor network structure; Motivations to be mentor characteristics; Effectiveness.	**McGinn et al. (2014)** 41 recent graduates of clinical research master's program.	**McGinn et al. (2014)** ***Content***: Developed new survey based on literature review. Refined survey based on pilot survey results (pilot survey methodology/results not reported). Survey focused on *mentor structure* (e.g., one mentor, multiple mentors, individual meetings, group meetings), *motivations to be mentored*, and overall *mentor effectiveness*. Complete list of relevant mentoring survey items provided. ***Internal Structure***: N/A – each indicator treated separately. ***Relationship with other mentoring relationship variables***: N/R.
Mentoring Triad Type [1] (Aikens et al., 2017; Aikens et al., 2016)	Mentoring Triad Type	**Aikens et al. (2016, 2017)** 842 undergraduates from 50 institutions across the U.S. All participants reported having completed one-or-more semesters (or summers) of research and were currently working with both a faculty and postgraduate on a research project.	**Aikens et al. (2016)** ***Content***: Developed single-item mentor triad type question (8 options for triad type) based on review of the literature and logical combinations of triad types. Complete list of relevant mentoring survey items provided. ***Internal Structure***: N/A – single item. ***Relationship with other mentoring relationship variables***: N/A. **Aikens et al. (2017)** ***Content***: N/A – see Aikens et al., (2016). ***Internal Structure***: N/A – see Aikens et al., (2016). ***Relationship with other mentoring relationship variables***: Moderate positive correlations between closed triad type, frequency of interaction with mentors, and rapport (i.e., psychosocial support).

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Mentorship Experience in College [24] (Gullan et al., 2016)	Challenge, Authenticity, Commitment, and Community	**Gullan et al. (2016)** 321 college students at a small private university (74% STEMM) (Gullan et al., 2016).	**Gullan et al. (2016)** ***Content***: Developed 66 items based on the work of Magolda (2009) – focusing on challenges provided by and support from mentors. In addition, items developed to assess duration and formality of the relationship. Complete list of relevant mentoring survey items provided. ***Internal Structure***: Used iterative principle components analyses to extract 4-components (i.e., parallel analysis, Kaiser rule, low communalities used to remove items and determine items to retain and components to extract). 24 of 66 items survived the iterative analysis. Components labeled Challenge, Authenticity, Commitment, and Community. Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: Small-to-moderate sized correlations among the four components.
Mentoring strategies and approaches [14] (Haeger & Fresquez, 2016)	Instrumental support, Socioemotional support, Culturally relevant support.	**Haeger & Fresquez (2016)** 138 undergraduate students at a public minority serving university (50% in science majors) (Haeger & Fresquez, 2016).	**Haeger & Fresquez (2016)** ***Content***: N/R. Complete list of relevant mentoring survey items provided. ***Internal Structure***: Used principle components analysis to determine 2-component structure, combined Socioemotional-Culturally relevant support and instrumental support. Despite this finding, authors chose to use original 3-factor approach. Internal consistency reliabilities reported at >0.70. ***Relationship with other mentoring relationship variables***: Moderate correlations among the three mentoring support factors.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Deaf Mentoring Survey [DMS, 15] (Braun, Gormally, & Clark, 2017)	Being a scientist, Deaf community capital, Asking for accommodations, and Communication access.	**Braun et al. (2017)** 59 deaf undergraduate students, graduate students, or postdocs in scientific disciplines.	**Braun et al. (2017)** ***Content***: Followed Thompson et al.'s (2015) framework to assess human, social, and cultural capital involved in mentoring relationships between mentors and deaf mentees. Theoretically derived seven aspects of capital. Adapted items from the literature to assess human (academic) capital. Developed new items based on focus group, cognitive interviews, and a pilot study with deaf individuals with research experiences. Total of 35 items developed to assess 6 of 7 types of capital. Complete list of relevant mentoring survey items provided. ***Internal Structure***: Used iterative exploratory factor analyses to remove items and determine the factor structures (Kaiser rule and scree plot used to determine factor extraction). Final EFA included 15 items and revealed 4 mentoring support factors labeled: Being a scientist, Deaf community capital, Asking for accommodations, and Communication access. These four factors included a mixture of capital types. Internal consistency reliabilities reported at >0.70. ***Relationship with other mentoring relationship variables***: N.R.
Evaluation of Mentoring Relationship [9] (Dennehy & Dasgupta, 2017)	Global measure of similarity, support and satisfaction.	**Dennehy & Dasgupta (2017)** 150 female first-year college students majoring in engineering at a public university.	**Dennehy & Dasgupta (2017)** ***Content***: Adapted items from prior research. Item contents contain a mixture of questions about perceived similarity with mentor, identification of mentor as a role model, and satisfaction with the relationship. Complete list of relevant mentoring survey items provided. ***Internal Structure***: Internal consistency reliabilities reported at >0.70. ***Relationship with other mentoring relationship variables***: N.A.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Developmental Network Mapping Activity [12] (Christou et al., 2017)	Network structure.	**Christou et al. (2017)** 107 academic medical faculty that took part in the Faculty Mentoring Leadership Program at Brigham and Women's Hospital.	**Christou et al. (2017)** ***Content***: Adapted the Kram/Higgins mentor network activity to academic medicine. Mentor network activity is a self-assessment on identifying current mentoring support from a mentor network in terms of 1) Diversity, 2) Redundancy, 3) Interconnectivity, 4) Strength, 5) Balance, and 6) Connections to power and influence. ***Internal Structure***: N.A. – self-assessment is a diagram rather than a scale. ***Relationship with other mentoring relationship variables***: N.A.
Scientific mentor network size [4] (Hernandez et al., 2017)	Network size.	**Hernandez et al. (2017)** 240 first- and second-year undergraduate women in STEMM majors (Hernandez et al., 2017).	**Hernandez et al. (2017)** ***Content***. Created a 4-item checklist of current mentorship received from university faculty, graduate students, peers, and scientific professionals outside the university. Questions derived from the types of scientific mentors represented in the mentoring in STEMM literature. ***Internal Structure***: N.R. ***Relationship with other mentoring relationship variables***:

Appendix 2.

Summary of Measures of Mentoring Used in STEMM from Mentor Perspective

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Mentor skills self-assessment [12] (Pfund et al., 2006)		**Pfund et al., (2006)** 52 mentors enrolled in mentor training seminar.	**Pfund et al., (2006)** ***Content***: Twelve indicators of mentor skill gains aligned with key elements of mentor training seminar: Address diversity; Assessing understanding; Building confidence; Build trust and respect; Communicate research; Dealing with challenges in mentoring; Establish expectations; Fostering independence; Giving feedback; Research planning. ***Internal Structure***: N/A – each indicator treated separately. ***Relationship with other mentoring relationship variables***: N/R.
Mentoring Competency Assessment [MCA, 26] (Fleming et al., 2013; Pfund et al., 2013; Pfund et al., 2014)	Maintaining effective communication, Aligning expectations, Assessing understanding, Addressing diversity, Fostering independence, & Promoting professional development.	**Fleming et al. (2013) & Pfund et al. (2014)** 283 mentor (faculty) –mentee (undergraduate through faculty) pairs from 16 universities involved in mentoring in STEMM training program.	**Fleming et al. (2013)** ***Content***: Aligned 26-indicators with 6-part framework and learning objectives of Mentor Training for Clinical and Translational Researchers workshop (Pfund et al., 2013). Pilot testing used cognitive interviews with mentors and mentees to assess the instrument's consistency. ***Internal Structure***: Confirmatory factor analysis conducted on the hypothesized 6-factor structure. Data-model fit indices showed less-than good-fit. Internal consistency reliability reported at >0.70 for Aligning expectations, Assessing understanding, Fostering independence, & Promoting professional development; but <0.70 for Maintaining effective communication and Addressing diversity. ***Relationship with other mentoring relationship variables***: Moderate to high correlations among the six factors.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Knowledge, Skills, and Attributes of Mentors [KSAM, 30] (Ahn & Cox, 2016)	Building personal relationship (*Psychosocial*); Building working relationship (*Psychosocial*); Meeting individual needs (*Instrumental*); & Daily research tasks (*Instrumental*).	**Ahn & Cox (2016)** 101 graduate students / postdocs that serve as mentors of undergraduate research assistants.	**Ahn & Cox (2016)** ***Content***: Collected data from 17 graduate student / postdoctoral mentors of undergraduate researchers using one-on-one interviews. Qualitative data analysis identified three themes where mentors provide support to mentees: knowledge (e.g., assess student knowledge/gaps), skills (research support skills), attributes (provide care) Researchers developed 57 items to align with the three themes. Research project mentors evaluated the items. Complete list of relevant mentoring survey items provided. ***Internal Structure***: Conducted multiple exploratory factor analyses to determine 4-factor solution. Used the Kaiser rule, scree plot, and factor loadings to determine factor extraction. Used iterative process to review 27 items based on factor analysis – either did not load on the hypothesized factor, loaded on multiple factors, or loaded on no factors. Resulting four factors measured: Building working relationship, Meeting individual needs, Daily research tasks, and Building personal relationship. Internal consistency reliability reported at >0.70. ***Relationship with other mentoring relationship variables***: N/R.

Appendix 3.

Summary of Measures of Mentoring Used in STEMM from Institutional and Program Evaluation Perspectives

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Institutional Efforts to Support Research Mentorship [12-69] (Keyser et al., 2008; Tillman et al., 2013)		**<u>Keyser et al. (2008)</u>** Proposed use of institutional self-assessment form to document institutional roles in supporting research mentorship. No empirical data (Keyser et al., 2008). **<u>Tillman et al. (2013) & Abedin et al. (2013)</u>** 51 Clinical and Translational Science Awarded institution (CTSA) and 53 KL2 education program leaders at host institutions (Abedin, Rebello, Richards, & Pincus, 2013; Tillman et al., 2013).	**<u>Keyser et al. (2008)</u>** ***Content***. Working group examined the literature and AAMC Compact between Postdoctoral Appointees and Their mentors to derive the following institutional mechanisms for monitoring research mentoring: "(1) the *criteria* for selecting mentors, (2) *incentives* for motivating faculty to serve effectively as mentors, (3) factors that facilitate the *mentor–mentee relationship*, (4) factors that strengthen a mentee's ability to conduct *research* responsibly, and (5) factors that contribute to the *professional development* of both mentees and mentors." ***Internal Structure***: N/A – no data collected. ***Relationship with other mentoring relationship variables***: N/A – no data collected. **<u>Tillman et al. (2013) & Abedin et al. (2013)</u>** ***Content***. Created a 69 item survey for CTSA and KL2 programs/institutions based on Keyser et al., (2008) five domain framework (*mentor criteria*, *mentor incentives*, *mentor–mentee relationships*, *mentor–mentee research*, and *mentee/mentor professional development*) for monitoring research mentoring. Survey focuses on *policies*, *activities*, and *structural* responsibilities for each of the five domains. ***Internal Structure***: N/A – each indicator treated separately. ***Relationship with other mentoring relationship variables***: N/R.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Institutional Incentives for Mentoring & Mentor Survey of Costs/Benefits of Mentoring [4-12] (Maisel et al., 2017)	Institutional incentives to mentor, Benefits of mentoring, & Costs of mentoring.	**Maisel et al. (2017)** 110 mentors named participating in the U.S. Department of Veterans Affairs Health Services Research and Development Service's mentored career development award (CDA) program.	**Maisel et al. (2017)** ***Content***. Created a four item survey of institutional incentives to mentor and a 12 item benefits/costs of mentoring scale for CDA mentors – influenced by the work of Keyser et al., (2008). ***Internal Structure***: Internal consistency reliability reported for benefit of mentoring at >0.70. ***Relationship with other mentoring relationship variables***: Positive correlation between institutional incentives to mentor and mentor perceived benefits of mentoring.
KL2 Program Mentoring Program Semi-Structured Interview [26] (Silet et al., 2010)		**Silet et al. (2010)** 46 KL2 program directors at CTSA institutions.	**Silet et al. (2010)** ***Content***. Conducted literature review of mentoring in general and in medical settings specifically to generate semi-structure interview focused on mentoring infrastructure for *mentor selection, communication of expectations,* formal evaluation of the mentoring relationship, and *mentor support and training*. ***Internal Structure***: N/A – each indicator treated separately. ***Relationship with other mentoring relationship variables***: N/R.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
KL2 Mentor Self-Evaluation tool [20] (Anderson, Silet, & Fleming, 2012)		**Anderson et al (2012)** Proposed use of institutional evaluation and mentor self-assessment template to document opportunities for development and discussions about challenges and opportunities. No empirical data.	**Anderson et al (2012)** ***Content***. Based on Silet et al. (2010), created framework for formally evaluation of mentoring in CTSAs and KL2. ***Internal Structure***: N/A – No empirical data. ***Relationship with other mentoring relationship variables***: N/A – No empirical data.
Mentor Development Program Graduate survey [26] (Feldman et al., 2009; Feldman et al., 2012)		**Feldman et al. (2012)** 38 faculty mentors that had graduated from the MDP program.	**Feldman et al. (2012)** ***Content***: Indicators of mentor skill gains, barriers, and resources needed appear to be aligned with key elements of MDP training. Skill gains focused on 1) Mentor skill (gains), 2) Barriers to mentoring, and 3) Resources needed to mentor. ***Internal Structure***: N/A – each indicator treated separately. ***Relationship with other mentoring relationship variables***: N/R.
Medical School Mentoring Program Design and Implementation Survey [45] **(Fornari et al., 2014)**		**Fornari et al. (2014)** 14 U.S. medical schools – from leadership in Offices of Student Affairs.	**Fornari et al. (2014)** ***Content***: Survey items created by the author(s) based literature review. Survey items focused on 1) mentor program goals, 2) program design elements and activities, 3) program formality, 4) incentives and barriers for program implementation, and 5) evaluation. Complete list of relevant mentoring survey items provided. ***Internal Structure***: N/A – each indicator treated separately. ***Relationship with other mentoring relationship variables***: N/R.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Faculty motivations to mentor checklist [26] (Morales, Grineski, & Collins, 2016)		**Morales et al. (2016)** 541 research active faculty members at 13 research intensive universities that host summer undergraduate researchers.	**Morales et al. (2016)** ***Content***: Developed 13 items to assess motivations to mentor based on Allen's (2007) framework including: 1) Organizational citizenship behavior, 2) Expected costs and benefits, 3) Situational factors, 4) Previous mentoring experience, and 5) demographic factors. Developed 13 items to assess incentives to mentor based on professional judgement and literature review. Complete list of relevant mentoring survey items provided. ***Internal Structure***: N/A – each indicator treated separately. ***Relationship with other mentoring relationship variables***: Dispositional factors, perceived benefits, and grant funding, were positively associated with willingness to mentor. While perceived costs and later career stage were negatively correlated with willingness to mentor. No associations among mentorship variables reported.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Mentor/Protégé Satisfaction Survey (23) (Martina, Mutrie, Ward, & Lewis, 2014)		**Martina et al. (2014)** 73 faculty mentors who completed the U.R. CTSI mentor training course and 59 of their mentees.	**Martina et al. (2014)** ***Content***: Developed mentoring relationship satisfaction survey based on the degree to which mentors/protégé were satisfied with the quality, usefulness, and total time spent supporting the protégé in seven areas: (1) teaching, (2) research, (3) clinical care, (4) presentation skills, (5) networking, (6) career development, and (7) work-life balance during the year of CTSI. Complete list of relevant mentoring survey items provided. ***Internal Structure***: N/A – each indicator treated separately from both mentor and protégé perspectives. ***Relationship with other mentoring relationship variables***: N.R.
Mentoring Communities Network Analysis [N/A] (Chariker et al., 2017)		**Chariker et al. (2017)** 57,831 doctoral students and their faculty mentors (402 Nobel laureates and 57,429 non-Nobel laureates) in chemistry, physics, physiology, and medicine, as listed in The Academic Family Tree (David, 2016).	**Chariker et al. (2017)** ***Content***: Drawing on the work of Zuckermen (1977), inferred that highly successfully scientists (Nobel laureate) should have higher numbers of similarly successful mentors and protégé in their academic family tree. Quantified mentoring relationship in terms of doctoral advising relationship listed in Academic Tree. ***Internal Structure***: N.A. ***Relationship with other mentoring relationship variables***: Positive correlation between Nobel laureate status and academic descendants Nobel laureate status – the advisees (1st, 2nd, etc. generation) more likely to win a Nobel prize.

Scale Name [#items] & (Original Reference)	Subscales	Evidence from STEMM context	Methodology and Validation Evidence in STEMM context
Mentor Skill Gains in Supporting Mentee Self-Efficacy [6] (Butz, Branchaw, Pfund, Byars-Winston, & Leverett, 2018)		**Butz et al. (2018)** 166 research mentors in STEMM disciplines at multiple institutions. Participating were taking part in a training program to promote research self-efficacy mentorship skills.	**Butz et al. (2018)** ***Content***: Developed 6 items to assess mentor skills to support mentee research self-efficacy based on Social Cognitive Career Theory. Retrospectively asked mentors to rate their skills to 1) defining sources of self-efficacy, 2) build mentee confidence in research, 3) employ strategies to build mentee confidence, 4) assessment mentee confidence, 5) recognize deficits in mentee confidence, and 6) assess the overall quality of the mentorship relationship. Complete list of relevant mentoring survey items provided. ***Internal Structure***: N/A – each indicator treated separately. ***Relationship with other mentoring relationship variables***: N.R.
Research Mentor Training Implementation Survey (Program Evaluation) [70] (Spencer et al., 2018)		**Spencer et al. (2018)** 281 participants of a facilitator training (FT) intending to implement a research mentor training (RMT). FT participants included faculty, training program directors, university administrators, and other (e.g., postdocs, instructors).	**Spencer et al. (2018)** ***Content***: Developed 70 items to assess learning gains (retrospective approach) aligned with the goals of the RMT facilitator training (i.e., 6-part framework; Pfund et al., 2013). The evaluation survey also included several open-ended items related to RMT implementation intentions and modifications. Complete list of relevant RMT training implementation survey items provided. ***Internal Structure***: N/A – each indicator treated separately. ***Relationship with other mentoring relationship variables***: N.R.

References

Abedin, Z., Biskup, E., Silet, K., Garbutt, J. M., Kroenke, K., Feldman, M. D., . . . Pincus, H. A. (2012). Deriving competencies for mentors of clinical and translational scholars. *Clinical and Translational Science, 5*(3), 273-280. doi:doi:10.1111/j.1752-8062.2011.00366.x

Abedin, Z., Rebello, T. J., Richards, B. F., & Pincus, H. A. (2013). Mentor training within academic health centers with clinical and translational science awards. *Clinical and Translational Science, 6*(5), 376-380. doi:10.1111/cts.12067

AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. Washington DC: American Educational Research Association.

Ahn, B., & Cox, M. F. (2016). Knowledge, skills, and attributes of graduate student and postdoctoral mentors in undergraduate research settings. *Journal of Engineering Education, 105*(4), 605-629. doi:doi:10.1002/jee.20129

Aikens, M. L., Robertson, M. M., Sadselia, S., Watkins, K., Evans, M., Runyon, C. R., . . . Dolan, E. L. (2017). Race and gender differences in undergraduate research mentoring structures and research outcomes. *CBE-Life Sciences Education, 16*(2). doi:10.1187/cbe.16-07-0211

Aikens, M. L., Sadselia, S., Watkins, K., Evans, M., Eby, L. T., & Dolan, E. L. (2016). A social capital perspective on the mentoring of undergraduate life science researchers: An empirical study of undergraduate–postgraduate–faculty triads. *CBE-Life Sciences Education, 15*(2). doi:10.1187/cbe.15-10-0208

Allen, T. D., Eby, L. T., O'Brien, K. E., & Lentz, E. (2008). The state of mentoring research: A qualitative review of current research methods and future research implications. *Journal of Vocational Behavior, 73*(3), 343-357. doi:https://doi.org/10.1016/j.jvb.2007.08.004

Anderson, L., Silet, K., & Fleming, M. (2012). Evaluating and giving feedback to mentors: New evidence-based approaches. *Clinical and Translational Science, 5*(1), 71-77. doi:10.1111/j.1752-8062.2011.00361.x

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. doi:Doi 10.1097/00001888-200501000-00017

Braun, D. C., Gormally, C., & Clark, M. D. (2017). The deaf mentoring survey: A community cultural wealth framework for measuring mentoring effectiveness with underrepresented students. *CBE Life Sciences Education, 16*(1), ar10. doi:10.1187/cbe.15-07-0155

Brown, D. E. (1991). *Human universals*: McGraw-Hill New York.

Butz, A., Branchaw, J., Pfund, C., Byars-Winston, A., & Leverett, P. (2018). Promoting STEM trainee research self-efficacy: A mentor training intervention. *Understanding Interventions, 9*(1), 3730.

Byars-Winston, A. M., Branchaw, J., Pfund, C., Leverett, P., & Newton, J. (2015). Culturally diverse undergraduate researchers' academic outcomes and perceptions of their research mentoring relationships. *International Journal of Science Education, 37*(15), 2533-2554.

Chang, M. J., Sharkness, J., Hurtado, S., & Newman, C. B. (2014). What matters in college for retaining aspiring scientists and engineers from underrepresented racial groups. *Journal of Research in Science Teaching, 51*(5), 555-580. doi:doi:10.1002/tea.21146

Chariker, J. H., Zhang, Y. H., Pani, J. R., & Rouchka, E. C. (2017). Identification of successful mentoring communities using network-based analysis of mentor-mentee relationships across Nobel laureates. *Scientometrics, 111*(3), 1733-1749. doi:10.1007/s11192-017-2364-4

Chen, Y., Watson, R., & Hilton, A. (2016). A review of mentorship measurement tools. *Nurse Education Today, 40,* 20-28. doi:https://doi.org/10.1016/j.nedt.2016.01.020

Christou, H., Dookeran, N., Haas, A., Di Frances, C., Emans, S. J., Milstein, M. E., . . . Seely, E. W. (2017). Establishing effective mentoring networks: Rationale and strategies. *MedEdPORTAL, 13*(10571). doi:https://doi.org/10.15766/mep_2374-8265.10571

Crisp, G. (2009). Conceptualization and initial validation of the college student mentoring scale (CSMS). *Journal of College Student Development, 50*(2), 177-194.

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. doi:10.1007/s11162-009-9130-2

Crisp, G., & Cruz, I. (2010). Confirmatory factor analysis of a measure of "mentoring" among undergraduate students attending a hispanic serving institution. *Journal of Hispanic Higher Education, 9*(3), 232-244. doi:10.1177/1538192710371982

Curtin, N., Malley, J., & Stewart, A. J. (2016). Mentoring the next generation of faculty: Supporting academic career aspirations among doctoral students. *Research in Higher Education, 57*(6), 714-738. doi:10.1007/s11162-015-9403-x

de Janasz, S. C., & Sullivan, S. E. (2004). Multiple mentoring in academe: Developing the professorial network. *Journal of Vocational Behavior, 64*(2), 263-283. doi:https://doi.org/10.1016/j.jvb.2002.07.001

Dennehy, T. C., & Dasgupta, N. (2017). Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proceedings of the National Academy of Sciences, 114*(23), 5964-5969. doi:10.1073/pnas.1613117114

Dilmore, T. C., Rubio, D. M., Cohen, E., Seltzer, D., Switzer, G. E., Bryce, C., . . . Kapoor, W. N. (2010). Psychometric properties of the mentor role instrument when used in an academic medicine setting. *Clinical and Translational Science, 3*(3), 104-108. doi:10.1111/j.1752-8062.2010.00196.x

Downing, R. A., Crosby, F. J., & Blake-Beard, S. (2005). The perceived importance of developmental relationships on women undergraduates' pursuit of science. *Psychology of Women Quarterly, 29*(4), 419-426. doi:10.1111/j.1471-6402.2005.00242.x

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. doi:10.1037//0021-9010.75.5.539

Eagan, M. K., Hurtado, S., Chang, M. J., Garcia, G. A., Herrera, F. A., & Garibay, J. C. (2013). Making a difference in science education: The impact of undergraduate research programs. *American Educational Research Journal, 50*(4), 683-713. doi:10.3102/0002831213482038

Eby, L., Butts, M., Lockwood, A., & Simon, S. A. (2004). Protégés negative mentoring experiences: Construct development and nomological validation. *Personnel Psychology, 57*(2), 411-447. doi:10.1111/j.1744-6570.2004.tb02496.x

Eby, L. T., Allen, T. D., Hoffman, B. J., Baranik, L. E., Sauer, J. B., Baldwin, S., . . . Evans, S. C. (2013). An interdisciplinary meta-analysis of the potential antecedents, correlates, and consequences of protégé perceptions of mentoring. *Psychological Bulletin, 139*(2), 441-476. doi:10.1037/a0029279

Eby, L. T., Durley, J. R., Evans, S. C., & Ragins, B. R. (2008). Mentors' perceptions of negative mentoring experiences: Scale development and nomological validation. *Journal of Applied Psychology, 93*(2), 358-373. doi:10.1037/0021-9010.93.2.358

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. doi:10.1006/jvbe.1996.1547

Feldman, M. D., Huang, L., Guglielmo, B. J., Jordan, R., Kahn, J., Creasman, J. M., . . . Brown, J. S. (2009). Training the next generation of research mentors: The university of California, San Francisco, clinical & translational science institute mentor development program. *Clinical and Translational Science, 2*(3), 216-221. doi:doi:10.1111/j.1752-8062.2009.00120.x

Feldman, M. D., Steinauer, J. E., Khalili, M., Huang, L., Kahn, J. S., Lee, K. A., . . . Brown, J. S. (2012). A mentor development program for clinical translational science faculty leads to sustained, improved confidence in mentoring skills. *Clinical and Translational Science, 5*(4), 362-367. doi:10.1111/j.1752-8062.2012.00419.x

Fiske, A. P. (1992). The four elementary forms of sociality: framework for a unified theory of social relations. *Psychological review, 99*(4), 689.

Fleming, M., House, S., Shewakramani, V., Yu, L., Garbutt, J., McGee, R., . . . Rubio, D. M. (2013). The mentoring competency assessment: Validation of a new instrument to evaluate skills of research mentors. *Academic Medicine : Journal of the Association of American Medical Colleges, 88*(7), 1002-1008. doi:10.1097/ACM.0b013e318295e298

Fornari, A., Murray, T. S., Menzin, A. W., Woo, V. A., Clifton, M., Lombardi, M., & Shelov, S. (2014). Mentoring program design and implementation in new medical schools. *Medical Education Online, 19*(1), 24570. doi:10.3402/meo.v19.24570

George, Y. S., & Neale, D. (2006). *Report from study group meeting to develop a research and action agenda on STEM career and workforce mentoring*. Retrieved from https://www.researchgate.net/profile/Yolanda_George/publication/238776791_Report_from_Study_Group_Meetings_to_Develop_a_Research_and_Action_Agenda_on_STEM_Career_and_Workforce_Mentoring/links/0a85e5318b27ccc3e0000000.pdf

Gershenfeld, S. (2014). A review of undergraduate mentoring programs. *Review of Educational Research, 84*(3), 365-391. doi:10.3102/0034654313520512

Ghosh, R. (2014). Antecedents of mentoring support: a meta-analysis of individual, relational, and structural or organizational factors. *Journal of Vocational Behavior, 84*(3), 367-384. doi:10.1016/j.jvb.2014.02.009

Glessmer, M. S., Wang, Y. V., & Kontak, R. (2012). Networking as a tool for earth science women to build community and succeed. *Eos, Transactions American Geophysical Union, 93*(41), 406-407. doi:10.1029/2012EO410011

Golde, C. M., & Dore, T. M. (2001). *At cross purposes: What the experiences of today's doctoral students reveal about doctoral education (www.phd-survey.org)*. Philadelphia, PA: A report prepared for The Pew Charitable Trusts.

Green, S. G. (1991). Professional entry and the adviser relationship: Socialization, commitment, and productivity. *Group & Organization Studies, 16*(4), 387-407. doi:10.1177/105960119101600404

Green, S. G., & Bauer, T. N. (1995). Supervisory mentoring by advisers: Relationships with doctoral student potential, productivity and commitment. *Personnel Psychology, 48*(3), 537-562. doi:10.1111/j.1744-6570.1995.tb01769.x

Griese, E. R., McMahon, T. R., & Kenyon, D. B. (2016). A research experience for American Indian undergraduates: Utilizing an actor–partner interdependence model to examine the student–mentor dyad. *Journal of Diversity in Higher Education*. doi:10.1037/a0040033

Gullan, R. L., Bauer, K., Korfiatis, P., DeOliveira, J., Blong, K., & Docherty, M. (2016). Development of a quantitative measure of the mentorship experience in college students. *Journal of College Student Development, 57*(8), 1049-1055. doi:10.1353/csd.2016.0099

Haeger, H., & Fresquez, C. (2016). Mentoring for inclusion: The impact of mentoring on undergraduate researchers in the sciences. *CBE Life Sciences Education, 15*(3), ar36. doi:10.1187/cbe.16-01-0016

Handelsman, J., Pfund, C., Lauffer, S. M., & Pribbenow, C. M. (2005). *Entering mentoring: A seminar to train a new generation of scientists*. Madison, WI: University of Wisconsin Press.

Hernandez, P. R., Bloodhart, B., Barnes, R. T., Adams, A. S., Clinton, S. M., Pollack, I., . . . Fischer, E. V. (2017). Promoting professional identity, motivation, and persistence: Benefits of an informal mentoring program for female undergraduate students. *PLOS ONE, 12*(11), e0187531. doi:10.1371/journal.pone.0187531

Hernandez, P. R., Estrada, M., Woodcock, A., & Schultz, P. W. (2016). Protégé perceptions of high mentorship quality depend on shared values more than on demographic match. *The Journal of Experimental Education, 85*(3), 450-468. doi:10.1080/00220973.2016.1246405

Hernandez, P. R., Hopkins, P. D., Masters, K., Holland, L., Mei, B. M., Richards-Babb, M., . . . Shook, N. J. (2018). Student integration into STEM careers and culture: A longitudinal examination of summer research mentors and project ownership. *CBE-Life Sciences Education, 17:ar50*(3), 1-14. doi:10.1187/cbe.18-02-0022

Higgins, M. C. (2000). The more, the merrier? Multiple developmental relationships and work satisfaction. *Journal of Management Development, 19*(4), 277-296. doi:10.1108/02621710010322634

Higgins, M. C., & Kram, K. E. (2001). Reconceptualizing mentoring at work: A developmental network perspective. *Academy of Management Review, 26*(2), 264-288. doi:10.5465/amr.2001.4378023

Higgins, M. C., & Thomas, D. A. (2001). Constellations and careers: Toward understanding the effects of multiple developmental relationships. *Journal of Organizational Behavior, 22*(3), 223-247. doi:10.1002/Job.66

Hoyt, C. L., Burnette, J. L., & Innella, A. N. (2012). I can do that: The impact of implicit theories on leadership role model effectiveness. *Personality and Social Psychology Bulletin, 38*(2), 257-268. doi:10.1177/0146167211427922

Huskins, W. C., Silet, K., Weber-Main, A. M., Begg, M. D., Jr, V. G. F., Hamilton, J., & Fleming, M. (2011). Identifying and aligning expectations in a mentoring relationship. *Clinical and Translational Science, 4*(6), 439-447. doi:10.1111/j.1752-8062.2011.00356.x

Jacobi, M. (1991). Mentoring and undergraduate academic success: A literature review. *Review of Educational Research, 61*(4), 505-532. doi:10.2307/1170575

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York, NY: Guilford Press.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: Guilford.

Keyser, D. J., Lakoski, J. M., Lara-Cinisomo, S., Schultz, D. J., Williams, V. L., Zellers, D. F., & Pincus, H. A. (2008). Advancing institutional efforts to support research mentorship: A conceptual framework and self-assessment tool. *Academic Medicine, 83*(3), 217-225. doi:10.1097/ACM.0b13e318163700a

Kram, K. E. (1985). *Mentoring at work: Developmental relationships in organizational life*. Lanham, MD, England: University Press of America.

La Guardia, J. G., Ryan, R. M., Couchman, C. E., & Deci, E. L. (2000). Within-person variation in security of attachment: A self-determination theory perspective on attachment, need fulfillment, and well-being. *Journal of Personality and Social Psychology, 79*(3), 367-384. doi:10.1037/0022-3514.79.3.367

Lee, L. S., Pusek, S. N., McCormack, W. T., Helitzer, D. L., Martina, C. A., Dozier, A. M., . . . Rubio, D. M. (2012). Clinical and translational scientist career success: Metrics for evaluation. *Clinical and Translational Science, 5*(5), 400-407. doi:10.1111/j.1752-8062.2012.00422.x

Lewis, V., Martina, C. A., McDermott, M. P., Trief, P. M., Goodman, S. R., Morse, G. D., . . . Ryan, R. M. (2016). A randomized controlled trial of mentoring interventions for underrepresented minorities. *Academic Medicine, 91*(7), 994-1001. doi:10.1097/acm.0000000000001056

Lockwood, P. (2006). "Someone like me can be successful": Do college students need same-sex gender role models? *Psychology of Women Quarterly, 30*(1), 36-46. doi:10.1111/j.1471-6402.2006.00260.x

Lockwood, P., & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology, 73*(1), 91-103. doi:10.1037/0022-3514.73.1.91

Maisel, N. C., Halvorson, M. A., Finney, J. W., Bi, X., Hayashi, K. P., Blonigen, D. M., . . . Cronkite, R. C. (2017). Institutional incentives for mentoring at the US Department of Veterans Affairs and universities: associations with mentors' perceptions and time spent mentoring. *Academic Medicine, 92*(4), 521-527.

Martina, C. A., Mutrie, A., Ward, D., & Lewis, V. (2014). A sustainable course in research mentoring. *Clinical and Translational Science, 7*(5), 413-419. doi:10.1111/cts.12176

McGinn, A. P., Lee, L. S., Baez, A., Zwanziger, J., Anderson, K. E., Seely, E. W., & Schoenbaum, E. (2015). Mentoring in clinical-translational research: A study of participants in master's degree programs. *Clinical and Translational Science, 8*(6), 746-753. doi:10.1111/cts.12343

McKenzie, J. F., Wood, M. L., Kotecki, J. E., Clark, J. K., & Brey, R. A. (1999). Establishing content validity: Using qualitative and quantitative steps. *American Journal of Health Behavior, 23*, 311-318.

Meagher, E., Taylor, L., Probsfield, J., & Fleming, M. (2011). Evaluating research mentors working in the area of clinical translational science: A review of the literature. *Clinical and Translational Science, 4*(5), 353-358. doi:10.1111/j.1752-8062.2011.00317.x

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749. doi:10.1037/0003-066X.50.9.741

Montgomery, B. L. (2017). Mapping a mentoring roadmap and developing a supportive network for strategic career advancement. *SAGE Open, 7*(2), 2158244017710288. doi:10.1177/2158244017710288

Morales, D. X., Grineski, S. E., & Collins, T. W. (2016). Influences on faculty willingness to mentor undergraduate students from another university as part of an interinstitutional research training program. *CBE Life Sciences Education, 15*(3), ar35. doi:10.1187/cbe.16-01-0039

Morales, D. X., Grineski, S. E., & Collins, T. W. (2018). Effects of gender concordance in mentoring relationships on summer research experience outcomes for undergraduate students. *Science Education, 0*(0), 1-22. doi:10.1002/sce.21455

Noe, R. A. (1988). An investigation of the determinants of successful assigned mentoring relationships. *Personnel Psychology, 41*(3), 457-479.

Noy, S., & Ray, R. (2012). Graduate students' perceptions of their advisors: Is there systematic disadvantage in mentorship? *Journal of Higher Education, 83*(6), 876-914. doi:10.1353/jhe.2012.0036

Packard, B. W.-L. (2003). Web-based mentoring: Challenging traditional models to increase women's access. *Mentoring and Tutoring, 11*(1), 53-65.

Packard , B. W. L., Walsh, L., & Seidenberg, S. (2004). Will that be one mentor or two? A cross-sectional study of women's mentoring during college. *Mentoring & Tutoring: Partnership in Learning, 12*(1), 71-85. doi:10.1080/1361126042000183039

Paglis, L., Green, S., & Bauer, T. (2006). Does adviser mentoring add value? A longitudinal study of mentoring and doctoral student outcomes. *Research in Higher Education, 47*(4), 451-476. doi:10.1007/s11162-005-9003-2

Peltz, C. M., & Raymond, D. M. (2016). Effects of associate degree nursing students' characteristics on perceptions and experiences of mentoring. *Journal of Nursing Education, 55*(5), 258-265. doi:http://dx.doi.org/10.3928/01484834-20160414-04

Pfund, C., Byars-Winston, A., Branchaw, J., Hurtado, S., & Eagan, K. (2016). Defining attributes and metrics of effective research mentoring relationships. *Aids and Behavior, 20*, S238-S248. doi:10.1007/s10461-016-1384-z

Pfund, C., House, S., Spencer, K., Asquith, P., Carney, P., Masters, K. S., . . . Fleming, M. (2013). A research mentor training curriculum for clinical and translational researchers. *Clinical and Translational Science, 6*(1), 26-33. doi:10.1111/cts.12009

Pfund, C., House, S. C., Asquith, P., Fleming, M. F., Buhr, K. A., Burnham, E. L., . . . Sorkness, C. A. (2014). Training mentors of clinical and translational research scholars: A randomized controlled trial. *Academic Medicine, 89*(5), 774-782. doi:10.1097/acm.0000000000000218

Pfund, C., Pribbenow, C. M., Branchaw, J., Lauffer, S. M., & Handelsman, J. (2006). The merits of training mentors. *Science, 311*(5760), 473-474. doi:10.1126/science.1123806

Prime, D. R., Bernstein, B. L., Wilkins, K. G., & Bekki, J. M. (2015). Measuring the advising alliance for female graduate students in science and engineering: An emerging structure. *Journal of Career Assessment, 23*(1), 64-78. doi:10.1177/1069072714523086

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

Rice, K. G., Choi, C.-C., Zhang, Y., Villegas, J., Ye, H. J., Anderson, D., . . . Bigler, M. (2009). International student perspectives on graduate advising relationships. *Journal of Counseling Psychology, 56*(3), 376-391. doi:10.1037/a0015905

Rorrer, A. S. (2016). An evaluation capacity building toolkit for principal investigators of undergraduate research experiences: A demonstration of transforming theory into practice. *Evaluation and Program Planning, 55*, 103-111. doi:https://doi.org/10.1016/j.evalprogplan.2015.12.006

Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research, 27*(2), 94-104.

Sadler, T. D., Burgin, S., McKinney, L., & Ponjuan, L. (2010). Learning science through research apprenticeships: A critical review of the literature. *Journal of Research in Science Teaching, 47*(3), 235-256. doi:10.1002/Tea.20326

Schlosser, L. Z., & Gelso, C. J. (2001). Measuring the working alliance in advisor–advisee relationships in graduate school. *Journal of Counseling Psychology, 48*(2), 157.

Schlosser, L. Z., & Gelso, C. J. (2005). The advisory working alliance inventory-advisor version: Scale development and validation. *Journal of Counseling Psychology, 52*(4), 650-654. doi:10.1037/0022-0167.52.4.650

Schunk, D. H., & Mullen, C. A. (2013). Toward a conceptual model of mentoring research: Integration with self-regulated learning. *Educational Psychology Review, 25*(3), 361-389. doi:10.1007/s10648-013-9233-3

Scott, J. (2017). *Social network analysis*. Thousand Oaks, CA: Sage.

Sharkness, J., DeAngelo, L., & Pryor, J. (2010). *CIRP construct technical report*. Retrieved from https://heri.ucla.edu/cirp-constructs/

Silet, K. A., Asquith, P., & Fleming, M. F. (2010). A national survey of mentoring programs for KL2 scholars. *Clinical and Translational Science, 3*(6), 299-304. doi:10.1111/j.1752-8062.2010.00237.x

Sorkness, C. A., Pfund, C., Asquith, P., & Drezner, M. K. (2013). Research mentor training: Initiatives of the university of wisconsin institute for clinical and translational research. *Clinical and Translational Science, 6*(4), 256-258. doi:doi:10.1111/cts.12085

Spencer, K. C., McDaniels, M., Utzerath, E., Rogers, J. G., Sorkness, C. A., Asquith, P., . . . Feldon, D. (2018). Building a sustainable national infrastructure to expand research mentor training. *CBE—Life Sciences Education, 17*(3), ar48. doi:10.1187/cbe.18-03-0034

Syed, M., Azmitia, M., & Cooper, C. R. (2011). Identity and academic success among underrepresented ethnic minorities: An interdisciplinary review and integration. *Journal of Social Issues, 67*(3), 442-468. doi:10.1111/j.1540-4560.2011.01709.x

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

Tillman, R. E., Jang, S., Abedin, Z., Richards, B. F., Spaeth-Rublee, B., & Pincus, H. A. (2013). Policies, activities, and structures supporting research mentoring: A national survey of

academic health centers with clinical and translational science awards. *Academic Medicine, 88*(1), 90-96. doi:10.1097/ACM.0b013e3182772b94

Vogt, D. S., King, D. W., & King, L. A. (2004). Focus groups in psychological assessment: enhancing content validity by consulting members of the target population. *Psychological Assessment, 16*(3), 231-243. doi:10.1037/1040-3590.16.3.231