Initial development of the faculty research self-efficacy scale (FaRSES): Evidence of reliability and validity

By: Kelly L. Wester, Laura Gonzalez, L. DiAnne Borders, and Terry Ackerman

Wester, K. L., Gonzalez, L., Borders, L. D., & Ackerman, T. (2019). Initial development of the faculty research self-efficacy scale (FaRSES): Evidence of reliability and validity. *Journal of the Professoriate, 10(2),* 78-99.

Abstract:

Research self-efficacy is important to measure among faculty, given its relationship with scholarly productivity. However, very few measures exist to assess research self-efficacy among faculty, and few of those have evidence of validity. Thus, the Faculty Research Self-Efficacy Scale (FaRSES) was developed, with initial validation occurring among a nationwide sample of social and behavioral science faculty. A second order factor model emerged, where dimensions on the first factor included both (a) General Research Process and Quantitative Research and (b) Qualitative Research. Both first order factors loaded on one general second order factor. Scores on the first and second order FaRSES had strong evidence of reliability and validity. Suggestions for future research are provided.

Keywords: research self-efficacy | faculty productivity | scale development

Article:

*****Note: Full text of article below**

Initial Development of the Faculty Research Self-Efficacy Scale (FaRSES): Evidence of Reliability and Validity

Kelly L. Wester
The University of North Carolina at Greensboro

Laura Gonzalez
The University of North Carolina at Greensboro

L. DiAnne Borders
The University of North Carolina at Greensboro

Terry Ackerman
University of Iowa

***Abstract**: Research self-efficacy is important to measure among faculty, given its relationship with scholarly productivity. However, very few measures exist to assess research self-efficacy among faculty, and few of those have evidence of validity. Thus, the Faculty Research Self-Efficacy Scale (FaRSES) was developed, with initial validation occurring among a nationwide sample of social and behavioral science faculty. A second order factor model emerged, where dimensions on the first factor included both (a) General Research Process and Quantitative Research and (b) Qualitative Research. Both first order factors loaded on one general second order factor. Scores on the first and second order FaRSES had*

Kelly L. Wester is a Professor at The University of North Carolina at Greensboro. **Laura Gonzalez** is an Associate Professor at The University of North Carolina at Greensboro. **L. DiAnne Borders** is a Professor at The University of North Carolina at Greensboro. **Terry Ackerman** is a Distinguished Visiting Professor at The University of Iowa.

strong evidence of reliability and validity. Suggestions for future research are provided.

Key words: research self-efficacy, faculty productivity, scale development

Introduction

Self-efficacy has been a popular and useful construct for decades, particularly due to its ability to motivate future behavior. Bandura (1997) defined self-efficacy as "beliefs in one's capability to organize and execute the courses of action required to produce given attainments" (p. 3). Self-efficacy typically refers to the belief about one's ability or skill to carry out a specific task successfully; a person can be efficacious in one task (e.g., teaching) but not in a different task (e.g., research). Research self-efficacy, then, would be the belief in one's ability to engage successfully in different components of the research process.

Consistently, research self-efficacy has been shown to relate directly to scholarly productivity, for both graduate students (e.g., Deemer, Martens, Haase, & Jome, 2009; Hemmings & Kay, 2016; Kahn & Scott, 1997; Pasupathy & Siwatu, 2014) and faculty (e.g., Hemmings & Kay, 2016; Pasupathy & Siwatu, 2014). Scholarly productivity is important for society, as research is known to build upon existing knowledge and advance a discipline (Walker, Golde, Jones, Bueschel, & Hutchings, 2008). Research is also important for graduate students, as a new skill to be mastered and a key component of degree completion. Specific to faculty, research is a main factor on which hiring, promotion, and tenure decisions are made (Potter, Higgins, & Gabbidon, 2011). Although research has mainly been considered an activity primarily for faculty at research institutions, this view is shifting, as demands to conduct research are being pressed on faculty at all types of institutions (Eagan & Carvey, 2015; Jacobs & Winslow, 2004; Lucas & Murry, 2002). According to the National Center for Education Statistics, approximately one-quarter of full time faculty report that 30% of their time at work is spent on research, making it one of the primary activities within the role (Cataldi, Bradburn, & Fahimi, 2005). Thus, the need for departments and institutions to enhance research self-efficacy among faculty is important so faculty can meet these increasing demands and expectations. However, the lack of research self-efficacy measures at the

faculty level leaves a gap in our knowledge of faculty research self-efficacy

While exploring research self-efficacy at the faculty level is important, the majority of the studies focused on research self-efficacy have been conducted at the graduate student level, where research skills are often being acquired and used for the first time. Therefore, most research self-efficacy measures have been created for graduate students specifically. These graduate student focused measures include the Research Self-Efficacy Scale (RSES; Bieschke et al., 1996), the Self-Efficacy in Research Measure (SERM; Phillips & Russell, 1994), and the Research Attitudes Measure (RAM; O'Brien, 1998). Each of these measures have been found to be psychometrically sound; however, each includes items specific to graduate students and makes it difficult to apply the measure to faculty level researchers. For example, the SERM (Phillips & Russell, 1994) contains items such as confidence in one's ability in "defending a thesis or dissertation," and the RAM (O'Brien et al., 1998 as cited in Forester, Kahn, & Hesson-McInnis, 2004) inquires about the confidence one has in his or her graduate level training. Although the RSES (Bieschke et al., 1996) does not contain items specific to graduate students, the RSES is more quantitatively focused (e.g., "perform experimental procedures," "interpret statistical outputs") and contains some items that may lack relevance for today's researcher (e.g., developing computer programs to analyze data, finding articles not in your library). While it was developed for doctoral students, the RSES has been used with faculty (e.g., Hager, St. Hill, Prunuske, Swanoski, Anderson, & Lutfiyya, 2016). But the use of the RSES was with only six faculty, making statistical interpretation of the scale impractical (Hager et al., 2016).

More recently, a few scales have been developed specifically for faculty. However, some limitations also exist within these faculty specific measures. One measure was created based on themes from a qualitative study with faculty at a Spanish University (Vera et al., 2011). This measure includes 13 items which assess a combination of self-efficacy related to teaching, research, and management. Most of the items are vague and tend to be double barreled, potentially assessing more than one construct. For example, items include "update the use of research methodologies from my own specialty even when it is difficult to me to gain access to them" and "train new researchers even when economic

means are insufficient." The content in these items is important to academic researchers; however, the wording makes it difficult to discern if the response is specific to research self-efficacy versus other conundrums, such as economic situations or access to information.

Another research self-efficacy scale designed for faculty is the 20-item measure created by Griffioen et al. (2013). However, Griffioen et al. (2013) created this measure specifically for their study, which included non-university higher education lecturers. Additionally, no psychometric properties regarding evidence of validity were provided. This is concerning, as some of the items included in the measure appear to be demographic in nature (e.g., years of experience doing research, educational level) or designed to assess scholarly productivity (e.g., involvement in research projects) as measured by other researchers (e.g., Kahn & Scott, 1997), rather than self-efficacy specifically.

Finally, Pasupathy and Siwatu (2011) created a measure of research self-efficacy specific to faculty. Their measure includes four subscales: general research self-efficacy, qualitative research self-efficacy, quantitative research self-efficacy, and mixed methods research self-efficacy. Although all participants take the general research self-efficacy subscale, for the other three subscales participants are routed to only the subscales related to the type of research they conduct. For example, participants who indicate they only conduct qualitative research are asked to take the general research self-efficacy subscale and the qualitative research self-efficacy subscale, but would not be provided the quantitative or mixed methods subscales. Although creating separate subscales makes sense given that self-efficacy is task specific (Bandura, 1997), this approach creates difficulty in comparisons across faculty as a whole. Additionally, separating specific types of research methodology and design fails to assist in determining where gaps in research self-efficacy or competency may be, given that research competency is a combination of breadth of knowledge and depth of skill related to specific research components and spans across various methodologies (Wester & Borders, 2014). Finally, no evidence of validity is currently available for this research self-efficacy measure.

Kahn and Scott (1997) originally hypothesized that research self-efficacy would have a direct effect on scholarly productivity, with most empirical research supporting this at the graduate student level across disciplines

(Deemer, Martins, Haase, & Jome, 2009; Hemmings & Kay, 2016; Kahn & Scott, 1997; Lambie & Vaccaro, 2011; Paglis, Green, & Bauer, 2006). Lechuga and Lechuga (2012) pointed to efficacy as important to understanding one's competence which in turn influences scholarly productivity. Therefore, it could be assumed that research self-efficacy among faculty would be an important contributing factor in faculty scholarly productivity as well. Therefore, the importance of having a reliable and valid measure of faculty research self-efficacy is evident.

In sum, given the importance of research self-efficacy among faculty, it is problematic that existing measures to assess faculty research self-efficacy are either too general, lack applicability to tenure track full time faculty, lack clarity of connection to the research process, result in the inability to compare efficacy across faculty members, and/or have minimal to no psychometric information available. Multiple psychometrically sound measures do exist at the graduate student level to measure research self-efficacy; however, most of these measures are either lengthy in nature, measure more about the research training environment than actual efficacy, lack the inclusion of relevant qualitative methodological procedures and analyses, or include items that may not be relevant to faculty researchers today. Therefore, the current researchers set out to develop and validate a new measure, the Faculty Research Self-Efficacy Scale (FaRSES), to address these gaps and assess research self-efficacy of faculty in academia. The overarching purpose was to explore the initial psychometric properties of the FaRSES across faculty in the social and behavioral science disciplines. The specific research questions explored in the current study were the following: What is the factor structure of the FaRSES? Are scores on the FaRSES reliable? Does the FaRSES have evidence of validity? Validity evidence was explored in relation to the RSES to test for convergent validity, and scholarly productivity.

Methods

Construction of the FaRSES

Before constructing items, faculty research self-efficacy was operationally defined as "The extent or strength of one's belief in one's own ability to complete tasks and reach goals related to research. This entails the entire research process from onset and idea inception to dissemination." This definition, and subsequent items, were developed

through understanding Bandura's (1997) description of self-efficacy along with a review of research competencies from various disciplines (e.g., American Academy of Health Behavior, 2005; Peterson, Peterson, Abrams, Stricker, & Ducheny, 2010; Richardson, 2006; Wester & Borders, 2014). Attention was paid to ensuring that both quantitative and qualitative methodologies and analyses were represented in the development of the FaRSES because both methodologies have been included in research competency lists (Wester & Borders, 2014). Finally, because this measure was specific to faculty research self-efficacy, two items were included regarding the ability to maintain a research agenda while also attending to other responsibilities within a faculty role, as well as the ability to maintain multiple research projects simultaneously. The ability to balance multiple roles and responsibilities has been noted as important by multiple researchers, especially for early career faculty (e.g., Eagan & Garvey, 2015; Magnuson et al., 2006; Milsom & Moran, 2015). The process of item creation resulted in a total of 21 items created for the FaRSES, with a 5-point Likert-type scale (1 = *disagree*, 2 = *somewhat disagree,* 3 = *neutral,* 4 = *somewhat agree,* and 5 = *agree*). Participants are first provided with the definition of research self-efficacy (noted above) and the instructions on the FaRSES: "Consider your ability to engage in various aspects of the research process. Take a moment and indicate the degree to which you agree you have the ability to engage successfully in the following research tasks," followed by the 21 items. All items on the FaRSES can be found in Table 1.

Table 1
Factor loadings for FaRSES based on maximum likelihood extraction (21 items)

Item #	Item wording	Factor 1	Factor 2	Item Mean	(*SD*)
1	Develop a research idea or question that will make a contribution to your field by addressing an important gap in the existing research literature	*.77*	-.13	5.44	(.72)
2	Design a study (e.g., selecting methods, procedures, sample, data collection tools) that will answer my research questions	*.84*	-.10	5.54	(.62)
3	Select the appropriate methodology to answer the research question	*.78*	-.10	5.54	(.62)
4	Access research participants via appropriate networks	*.71*	.00	4.89	(1.19)
5	Obtain participant consent for my study (e.g., adult participants, minor participants, parent/guardian participants)	*.64*	-.05	5.33	(.85)
6	Plan and implement procedures for the study, which includes sample selection and various aspects of data collection (e.g., controlling extraneous variables, removing researcher bias, organizing interview questions, developing survey packet)	*.82*	-.15	5.39	(.78)
7	Implement a study utilizing quantitative methodology (e.g., experimental designs, correlational designs, longitudinal studies)	*.57*	-.30	5.28	(1.10)
8	Select instrumentation that validly and reliably measures my construct(s)	*.61*	-.25	5.30	(.88)
9	Implement a study utilizing qualitative methodology (e.g., phenomenology, ethnography, content analysis, grounded theory)	.18	*.93*	4.52	(1.54)
10	Utilize inferential and descriptive statistical analyses to analyze data	*.54*	-.16	5.23	(1.08)
11	Employ appropriate methods to analyze qualitative data	.20	*.86*	4.60	(1.55)
12	Seek out resources to support research agenda when needed (e.g., trainings, mentors, collaborators, funding sources, literature)	*.58*	.09	4.95	(1.13)
13	Identify appropriate professional journals or outlets to disseminate results through written text (e.g. manuscripts, reports)	*.73*	-.09	5.42	(.80)
14	When writing, use clear language and logical reasoning to introduce my research question in an introduction, and then move to appropriate technical language to describe the methods, and results	*.61*	-.14	5.56	(.60)
15	The ability to succinctly synthesize meaning from the results to develop implications	*.65*	-.11	5.48	(.65)
16	Accept and respond effectively to written or verbal constructive criticism of my research and writing	*.58*	.01	5.44	(.70)
17	Present findings to other professionals (e.g., conference proceedings, community partners)	*.60*	-.04	5.61	(.64)
18	Balance multiple research projects at various stages simultaneously	*.75*	-.07	5.06	(1.02)
19	Make sufficient progress on research projects while also attending to other job responsibilities (e.g., teaching/service)	*.63*	.02	4.21	(1.39)
20	Model appropriate research conduct to students involved in my research (e.g., observing, participating)	*.69*	-.08	5.15	(.97)
21	Engage in the entire research process, from start to finish (e.g., conceptualize, design, collect, analyze, and disseminate)	*.79*	-.04	5.39	(.95)

Note: Italicized factor loadings were retained on that factor.

Participants

A nationwide sample of faculty in the social and behavioral sciences employed at research universities were targeted for the current study. A total of 196 faculty opened and responded to the survey; however, 58 faculty were removed due to more than 50% missing data on the survey questionnaire, leaving a final sample of 138 faculty for this study. Over half of the respondents identified as female (61.6%), with 38.4% male. The majority (88.4%) self-identified as Caucasian, with others indicating African American (6.5%), Asian/Asian American (1.4%), Hispanic/Latino (1.4%), multiracial (1.4%), or other (.7%) as their racial or ethnic identity. Faculty participants represented a range of academic fields: psychology (n = 26, 18.8%), criminal justice/justice studies (n = 22, 15.9%), sociology (n = 15, 10.9%), family studies/human development (n = 14, 10.1%), public health (n = 14, 10.1%), social work (n = 11, 8%), counselor education (n = 6, 4.3%), women/gender studies (n = 3, 2.2%), and other (n = 27, 19.6%). Their faculty ranks included tenure track assistant professors (37.5%), associate faculty with tenure (25.4%), full professors with tenure (29%), associate clinical/non-tenure (8.7%), full clinical or non-tenure (2.9%), and fulltime instructor non-tenure (2.9%) (one participant did not respond to this demographic item). The university Carnegie classifications of respondents were doctoral research university (n = 53, 38.4%), high research university (n = 33, 23.9%), and very high research university (n = 43, 31.2%; 9 participants did not respond, 6.5%).

Measures

Faculty participants completed a demographic survey, the FaRSES, the RSES (Bieschke et al., 1996), and a measure of scholarly productivity. All measures were included to assist in providing evidence of construct or convergent validity for the FaRSES.

Faculty Research Self-Efficacy Scale. Information about the FaRSES is noted above in the construction of the scale.

Research Self-Efficacy Scale. The RSES was developed by Bieschke et al. (1996) to measure graduate students' estimation of their ability to perform various research related behaviors; however, it has been used with faculty (Hager et al., 2016). The RSES was tested among doctoral

students from various disciplines, including biological sciences, social sciences, physical sciences, and humanities. Evidence of internal consistency was found using Cronbach alpha of .96 for the total 51-item RSES. In addition, there was adequate evidence of construct validity, with expected correlations with research interest, number of years in graduate school, and involvement in research activities.

Demographic survey and scholarly productivity. Faculty were asked to indicate their gender, race, faculty rank, discipline, and university Carnegie classification. Scholarly productivity was measured by asking participants to respond to six items based on Kahn and Scott's (1997) Past Research Productivity and Current Research Involvement scale. Specifically, faculty were asked to report total publications published throughout their career, number of articles they had submitted but were not yet reviewed, manuscripts currently in preparation, funded research projects, and scholarly presentations at local, regional, and national conferences. Each individual item on the scholarly productivity scale could be utilized by itself. However, a sum score of productivity was calculated by adding the number of total publications, number of funded research projects, and the number of articles currently submitted for review. The number of manuscripts currently in preparation was not calculated into the total of scholarly productivity, as they were not completed scholarly products. Presentations also were not included in the total of scholarly productivity. Even though they could be deemed as scholarly products given the nature of disseminating findings to a larger audience, presentations at conferences vary widely based on discipline, with some conferences requiring working papers and others 15 minute presentations.

Procedures

Using the Carnegie classification database, half of the universities in each of the three Carnegie classifications of doctoral research university (n = 45), high research university (n = 21), and very high research university (n = 54) were randomly selected. We then searched the websites of each university to locate departments representing the specified social and behavioral sciences. Following Institutional Review Board (IRB) approval, the department chair for each department at the selected institutions was identified on the departmental website. Department chairs were sent an email requesting that the email and study

information be forwarded to each faculty member within their department, inviting participation. Follow-up emails were sent to department chairs two weeks later. The email invitation included a link to the consent document and the online survey. Specifically, we emailed 152 department chairs at 40 universities classified as Doctoral Research University (five universities had no contact or departmental information online), 106 department chairs at 19 universities classified as Research High University (two universities had no contact information online), and 276 department chairs at 54 universities classified as Research Very High University.

Data Analysis

Basic descriptive statistics and reliability of the scores were calculated for all measures. Exploratory factor analysis (EFA) was used to investigate the factor structure of the FaRSES. Given the nature of the EFA, maximum likelihood extraction method was used as suggested by Watson (2017). Correlations were conducted to examine relationships with other measures to provide evidence of construct and convergent validity. Finally, given increasing research expectations among faculty at all universities, exploration of the FaRSES across faculty rank and Carnegie classification were explored to better understand how research self-efficacy may be influenced by these categorizations.

Results

Prior to exploring the factorability of the FaRSES, basic means and standard deviations were computed for all measures. For scholarly productivity, faculty reported an average of 54.96 scholarly products (*SD* = 68.60). Breaking down individual components of scholarly productivity, faculty reported an average of 36.20 published articles (*SD* = 50.97), with 11.13 manuscripts currently submitted for review but not yet published (*SD* = 28.00). Faculty indicated being involved in an average of 4.92 funded activities (*SD* = 9.00). For total publications, which included published articles, submitted manuscripts, and funded projects, the mean was 44.09 (*SD* = 52.89, range 1 to 272; note that one person was removed as they reported having 1,000,000 total scholarly products, which was considered an outlier).

FaRSES EFA

In order to explore the factor structure of the FaRSES, we first examined the sphericity and sampling adequacy in the sample to ensure the data were factorable. Bartlett's test of sphericity was statistically significant (χ^2= 1573.13, *df* = 210, *p* <.001) and the Kaiser-Meyer-Olkin KMO (Kaiser, 1970) measure of sampling adequacy was adequate for factoring (.88). Kaiser (1970) noted that a KMO above .80 was deemed as "meritourious." Thus, based on these findings, an EFA was conducted.

Eigenvalues suggested the possibility of four factors (Factor 1 eigenvalue = 9.44, variance explained 44.95%; Factor 2 eigenvalue = 2.19, 10.41% variance explained; Factor 3 eigenvalue = 1.47, 7.00% variance explained; Factor 4 eigenvalue = 1.35, 6.44% variance explained). The scree plot supported a one-factor solution, while showing a second elbow at factor 3, suggesting the possibility of a two-factor solution (see Figure 1). Therefore, the possibility of a one-factor as well as a two-factor solution was explored (see Table 1). Based on the factor loadings, all items loaded on Factor 1 (above .50 for each item) with the exception of the two qualitative methodology and analysis items (items 9 and 11), which loaded high on Factor 2 (above .85 each). Inter-item correlations can be found in Table 2.

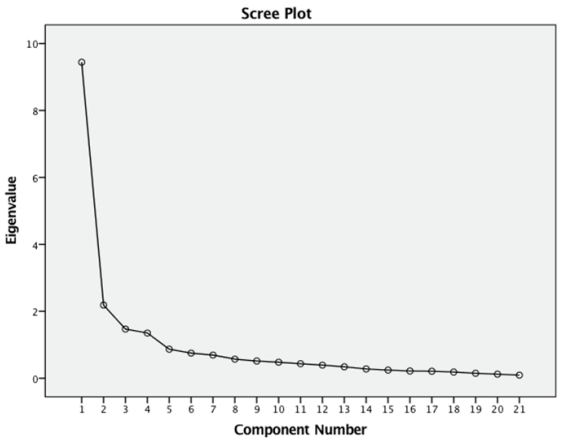

Figure 1. Scree plot from FaRSES EFA, 21 Items

Table 2
Correlations between 21 individual items on the FaRSES

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	-	.75**	.64**	.57**	.39**	.57**	.37**	.43**	.06	.40**	.12	.46**	.62**	.43**	.51**	.35**	.43**	.56**	.49**	.52**	.55**
2		-	.75**	.49**	.37**	.65**	.42**	.39**	.11	.41**	.12	.56**	.63**	.51**	.58**	.45**	.57**	.57**	.42**	.53**	.55**
3			-	.42**	.33**	.57**	.40**	.37**	.04	.41**	.09	.39**	.61**	.48**	.58**	.49**	.46**	.48**	.39**	.51**	.49**
4				-	.68**	.66**	.36**	.32**	.19*	.33**	.24*	.52**	.59**	.33**	.36**	.37**	.29**	.59**	.59**	.65**	.54**
5					-	.63**	.29**	.15	.12	.25**	.09	.39**	.47**	.21*	.29**	.26**	.36**	.41**	.48**	.58**	.46**
6						-	.47**	.42**	.08	.41**	.10	.48**	.59**	.48**	.58**	.41**	.50**	.57**	.42**	.60**	.54**
7							-	.69**	-.14	.78**	-.09	.32**	.46**	.38**	.39**	.19*	.29**	.42**	.34**	.49**	.36**
8								-	-.08	.60**	-.00	.21*	.45**	.42**	.48**	.25**	.37**	.42**	.32**	.41**	.34**
9									-	.04	.87**	.28**	.08	.01	.04	.18	.12	.15	.18	.13	.15
10										-	.16	.39**	.41**	.31**	.41*	.17	.34**	.46**	.49**	.46**	.34**
11											-	.32**	.12	.02	.12	.26**	.19*	.20*	.28**	.14	.17
12												-	.63**	.22*	.22*	.40**	.43**	.50**	.50**	.49**	.42**
13													-	.48**	.42**	.43**	.48**	.55**	.54**	.60**	.55**
14														-	.67**	.45**	.39**	.39**	.35**	.39**	.33**
15															-	.51**	.47**	.48**	.32**	.37**	.48**
16																-	.41**	.40**	.33**	.30**	.29**
17																	-	.43**	.29**	.36**	.39**
18																		-	.64**	.64**	.61**
19																			-	.60**	.52**
20																				-	.68**
21																					-

Exploring the internal consistency of the FaRSES, there was evidence of consistency of responses to the items, with a Cronbach alpha of .91 on the one-factor solution (i.e., all 21 items on one factor). Given the previously noted factor loadings (found in Table 1), reliability was explored within the two-factor solution as well, finding high evidence of reliability for both (subscale 1 (19 items) $\alpha = .94$; subscale 2 (2 items) $\alpha = .92$).

Given that items 9 and 11 did not load on Factor 1, and were poorly correlated with other items on factor 1, we further explored whether these two items needed to be removed from the larger scale, or whether another factor solution existed. Specifically, it was found that the two potential subscales were not correlated ($r = .09$, $p > .05$); however, they both highly loaded on the larger single factor solution of faculty research self-efficacy (FaRSES subscale 1 and FaRSES subscale 2 loaded on the total FaRSES with factor loadings of .74 each, explaining a total of 54.48% of the variance, $\alpha = .91$). Therefore, it was determined that this was a second-order factor solution, with the first order factors being FaRSES Factor 1 (entitled Self-Efficacy of General Research Process and Quantitative Methodology, 19 items) and FaRSES Factor 2 (entitled Self-Efficacy of Qualitative Methodology, 2 items), and the factor loadings of both of these factors loading onto the second order factor of Faculty Research Self-Efficacy (Total FaRSES).

Before moving forward to test for evidence of validity, this solution was further tested using hierarchical cluster analysis (Roussos, Stout, & Marden, 1998). The results from the cluster analysis supported the two-factor solution based on the agglomeration schedule coefficients. The dendrogram (see Figure 2) provides the linkage between clusters and the point in which the items joined. As can be seen in the dendrogram for the FaRSES items, items 9 and 11 (which assess efficacy in qualitative methodology and analysis) stand in their own cluster, with item 9 joining the other cluster of items at the final stage. Given the results from the EFA and the cluster analysis, the FaRSES is considered a second order factor model, with first order factors of (1) Efficacy of General Research Process and Quantitative Research and (2) Efficacy of Qualitative Research, with a second order factor of Faculty Research Self Efficacy.

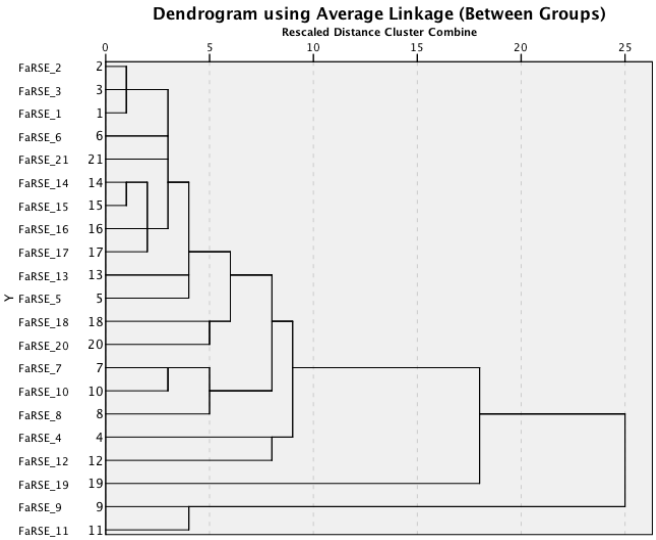

Figure 2. Dendogram of 21 FaRSES items

Evidence of Validity

To test for evidence of construct validity, the first order and second order FaRSES factors were correlated with Bieschke's research self-efficacy scale (RSES), as well as total scholarly productivity. The second order FaRSES factor correlated with the RSES at a moderately high level (r = .63, p < .01). When exploring the first order FaRSES factors, only Factor 1 (Efficacy of General Research Process and Quantitative Research) correlated at a moderate level with the RSES (r = .64, p < .01), whereas the first order Factor 2 (Efficacy of Qualitative Research) was not significantly correlated with the RSES (r = .08, p > .05). The lack of statistically significant correlation of Factor 2 and the RSES makes sense, given the lack of qualitative items on Bieschke et al.'s 51 item RSES.

In regard to total scholarly productivity, the second order FaRSES factor (total score) was positively related to total scholarly productivity at a moderate to high level (r = .43, p < .01), with moderate to high correlations for each type of scholarly productivity. More specifically, the second order FaRSES factor was positively correlated with the number of published articles (r = .43, p < .01), number of articles

submitted for review (r =.19, p <.05), but not to the number of funded projects (r = .11, p > .05). Regarding the first order FaRSES factors, Factor 1 was positively and statistically related to total scholarly productivity (r = .46, p < .05); however, Factor 2 was not statistically related (r = -.05, p >.05).

Given the recent suggestions that research is carried out at all institutions (Eagan & Carvey, 2015; Jacobs & Winslow, 2004; Lucas & Murry, 2002), differences across Carnegie research classifications were explored (i.e., doctoral research, research high, research very high universities). Significant differences were found on the second order FaRSES factor (F(2, 118) = 4.31, p < .05) and the first order FaRSES Factor 1 (F(2, 122) = 5.42, p < .01) across classifications, but the first order FaRSES Factor 2 did not significantly differ across classifications (F(2, 121) = .47, p > .05). Exploring post-hoc Schéffe tests to further understand the significant differences across Carnegie classifications on the second order FaRSES (full scale score) and the first order FaRSES Factor 1, faculty from doctoral research university classifications did not significantly differ on the second order FaRSES Factor (M = 109.74, SD = 12.08) from faculty at Research High (M = 103.38, SD = 13.63) or Research Very High classifications (M = 111.92, SD = 12.11). However, faculty from Research High did significantly differ from faculty at Research Very High classifications. In exploring the first order Factor 1 (Efficacy in General Research Process and Quantitative Research) similar findings existed, with faculty at Research High (M = 94.29, SD = 12.53) significantly differing from faculty at Research Very High (M = 103.33, SD = 11.76) classifications, with no statistical difference among faculty at Doctoral Research University (M = 100.52, SD = 11.17) classifications from the other two classifications.

In exploring differences in research self-efficacy among faculty across rank, significant differences were found on the second order FaRSES factor (F(2, 106) = 12.58, p < .001) and the first order FaRSES Factor 1 (F(2, 112) = 13.33, p < .001) but not on the first order FaRSES Factor 2 (F(2, 111) = .56, p < .05). Only tenure track or tenured faculty were explored in this analysis due to the small samples sizes of clinical faculty categories. For the second order FaRSES factor, full professors (M = 109.32, SD = 12.49) reported having higher levels of self-efficacy than did both assistant and associate professors (M = 105.26, SD = 12.96; M = 105.09, SD = 12.54, respectively), which was similar to differences

found on the first order FaRSES Factor 1. No significant differences were found between assistant and associate professors on the second order FaRSES Factor or on the first order FaRSES Factor 1. No differences were found across rank on the first order FaRSES Factor 2.

Discussion

The purpose of the current study was to pilot and provide initial evidence of validity and reliability for a new measure of faculty research self-efficacy, the FaRSES. Based on the results of a national sample of social and behavioral science faculty, primarily tenured or on the tenure track, it appears the FaRSES has potential as a measure of faculty research self-efficacy. Although the sample size was relatively small, the initial examination of the FaRSES revealed a second order factor solution, with two first order factors. Evidence of internal consistency, given the strong Cronbach alpha scores, as well as support for construct validity, given the correlations with Bieschke et al.'s (1996) RSES measure and scholarly productivity indicators, suggest that the FaRSES in fact does measure faculty research self-efficacy.

Results from the current study suggested the possibility of two first order factors that measure separate constructs, but that load together on one second order factor of research self-efficacy. Specifically, the two qualitative items consistently did not factor onto the first primary factor, whereas all other items did–including the quantitative methodology and statistical analysis items. Although the two qualitative items did not specifically load as items onto the scale, when they were combined into a first order factor of qualitative research self-efficacy specifically, this first order factor did load as one of two first order factors onto a single second order factor of the FaRSES. However, the consistent inability of the two qualitative methodological items to load onto one larger main factor of research self-efficacy with all other items is puzzling.

The failure of the qualitative items to load on the same factor with the other 19 items assessing for research process and quantitative methodology and analysis may be why other researchers, (i.e., Pasupathy & Siwatu, 2011), have separated these items into completely different scales, although the reason they kept their quantitative and qualitative scales completely independent is not clearly stated in their discussion of their research self-efficacy measure. However, considering the larger

latent construct of research self-efficacy, the exclusion of qualitative items from an overall research self-efficacy scale does not allow the ability to compare across faculty who may be engaging in different types of qualitative or quantitative research due to philosophical or methodological beliefs or norms about research. Additionally, various research competency lists suggest that knowledge and skill in both qualitative and quantitative methodologies are important for overall research competency (e.g., Wester & Borders, 2014); thus, it seems that researchers need self-efficacy in both – or at minimum the ability to consider both qualitative and quantitative methodologies when considering overall research efficacy.

In exploring the correlations in an attempt to explain the lack of connection of the two qualitative items on the FaRSES, it appears that the qualitative items had low or negative correlations with quantitative items, as well as most other items on the FaRSES, with the exception of the statistically positive relationship with 'access to research participants' (item 4) and 'modeling appropriate research conduct to students involved in my research' (item 20). It should also be noted that, when exploring the distribution of responses on each item, the two qualitative items had wider ranges in responses, with more faculty indicating lack of agreement in their efficacy on qualitative methodology and analysis than on other items on the FaRSES. A brief exploration into discipline was conducted by examining mean scores of the subscales; results suggested that some differences by discipline may exist among the qualitative items, with some social and behavioral sciences faculty overall reporting less efficacy in the Efficacy of Qualitative Research first order FaRSES Factor 2 than the Efficacy of the General Research Process and Quantitative Research first order FaRSES Factor 1. A formal statistical analysis was not conducted to explore these mean score differences by discipline due to the small sample sizes per discipline, which would likely result in lack of statistical power to find meaningful results. However, the efficacy in qualitative research versus quantitative research and other components of the research process by discipline should be given further consideration, as the differences across disciplines regarding efficacy of qualitative research may link to philosophical or paradigmatic beliefs within disciplines which could impact how faculty were trained in research.

The lack of significant relationships between the first order FaRSES Factor 2 (Efficacy of Qualitative Research) in relation to scholarly productivity and RSES is important to note. Although reliability analysis for this specific subscale is adequate, it does draw into question the ability of this subscale to truly measure efficacy of research in qualitative methodologies in isolation. Two items, ultimately, cannot asses the depth and breadth of efficacy in qualitative methodology. Depth and breadth of knowledge in specific research tasks related to qualitative designs are important (Wester & Borders, 2014). Additionally, Watson (2017) summarized scholars who suggested the importance of having three to ten items to measure a construct. While the original goal in developing the FaRSES was not to have two separate subscales, but one overarching scale that assessed the breadth of faculty research self-efficacy, the resultant second order factor structure emerged. Thus, given that qualitative methodology is on a first order factor independently, it may be important to add additional qualitative methodological and analytical items to first order FaRSES Factor 2 for it to function adequately. However, it should be noted that first order FaRSES Factor 2 does strongly load onto the second order FaRSES factor with the first order FaRSES Factor 1 (Efficacy of General Research Process and Quantitative Research). Thus, the current authors recommend using the second order FaRSES factor, or the first order FaRSES Factor 1, but currently do not recommend the use of the first order FaRSES Factor 2 in isolation.

Limitations and Directions for Future Research

Our study has several limitations worth noting. First, even though our sample was a randomly generated sample of social and behavioral science departments within research classified schools (as noted by Carnegie classification), the list of social and behavioral science disciplines may not have been exhausted, nor was information available for each department on their university website. This may have led to departments, and thus faculty, being unintentionally left out of the sampling. Additionally, although this sample was a nationwide sample, the resulting sample size was small, limiting our ability to conduct a confirmatory factor analysis, as well as limiting generalizability. As can be seen from the results, it appears that qualitative methodology may be important to some disciplines, but potentially less important or less of a focus in others. Therefore, future researchers using the FaRSES need to

increase sample sizes within each discipline to determine the stability of the factor loadings, exploring between and, potentially, also within discipline self-efficacy. Additionally, as noted, two items on one first order factor may not be sufficient to measure an entire construct. Thus, additional qualitative items need to be added to the first order factor to measure the depth and breadth of qualitative research if Factor 2 is to be used in isolation. Otherwise the second order FaRSES factor should be used in its entirety.

Although more research needs to be conducted to explore the possibility of why the qualitative items did not load on the same factor as the other FaRSES items, it should be noted that ultimately all 21 items remained on the second order FaRSES factor solution, with two first order factors.

Conclusion

The FaRSES was specifically designed for faculty researchers, as it inquires about aspects of the research process, quantitative and qualitative methodologies and analysis, as well as the ability to balance research with other faculty responsibilities. Faculty could contemplate use of the FaRSES as a predictive tool regarding scholarly productivity or as a developmental tool to examine faculty research self-efficacy. Research self-efficacy, and the ability to measure it among faculty, becomes continually more important to faculty success in academia.

References

American Academy of Health Behavior (2005). A vision for doctoral research training in health behavior: A position paper from the American Academy of Health Behavior. *American Journal of Health Behavior, 29,* 542-556.

Bandura (1997). *Self-efficacy: The exercise of control.* New York, NY: Freeman.

Bieschke, K. J., Bishop, R. M., & Garcia, V. L. (1996). The utility of the Research Self-efficacy Scale. *Journal of Career Assessment, 4,* 59-75.

Cataldi, E. F., Bradburn, E. M., & Fahimi, M. (2005). 2004 National study of postsecondary faculty (NSOPF:04): Background characteristics, work activities, and compensation of instructional faculty and staff, Fall 2003. E.D. TAB. NCES 2006-176. Retrieved online: https://eric.ed.gov/?id=ED489118

Deemer, E. D., Martens, M. P., Haase, R. F., & Jome, L. M. (2009). Do mastery approach goals mediate the relationship between the research training environment and research interest? Test of a social cognitive model. *Training and Education in Professional Psychology, 3,* 250-260.

Eagan, M. K., & Garvey, J. C. (2015). Stressing out: Connecting race, gender, and stress with faculty productivity. *The Journal of Higher Education, 86,* 923-954.

Forester, M., Kahn, J. H., & Hesson-McInnis, M. S. (2004). Factor structures of three measures of research self-efficacy. *Journal of Career Assessment, 12,* 3-16.

Griffioen, D. M. E., de Jong U., & Jak, S. (2013). Research self-efficacy of lecturers in non-university higher education. *Innovations in Education and Teaching International, 50,* 25-37.

Hager, K., St Hill, C., Prunuske, J., Swanoski, M., Anderson, G., & Lutfiyya, M. N. (2016). Development of an interprofessional and interdisciplinary collaborative research practice for clinical faculty. *Journal of Interprofessional Care, 30,* 265-267.

Hemmings, B., & Kay, R. (2016). The relationship between research self-efficacy, research disposition, and publication outlet. *Educational Psychology, 36,* 347-361.

Jacobs, J. A., & Winslow, S. E. (2004). Overworked faculty: Job stresses and family demands. *ANNALS AAPSS, 596,* 104-129.

Kahn, J. H., & Scott, N. A. (1997). Predictors of research productivity and science related career goals among counseling psychology graduate students. *The Counseling Psychologist, 25,* 38-67.

Kaiser, H. F. (1970) A second generation little jiffy. *Psychometrika, 35(4),* 401–415. Lambie, G. W., & Vaccaro, N. (2011). Doctoral counselor education students' levels of research self-efficacy, perceptions of research training environment, and interest in research. *Counselor Education and Supervision, 50,* 243-258.

Lechuga, V. M., & Lechuga, D. C. (2012). Faculty motivation and scholarly work: Self-determination and self-regulation perspectives. *Journal of the Professoriate, 6,* 59-97.

Lucas, C. J., & Murry, J. W., Jr. (2002). *New faculty: A practical guide for academic beginnings.* New York, NY: Palgrave.

Magnuson, S., Black, L. L., & Lahman, M. K. E. (2006). The 2000 cohort of new assistant professors of counselor education: Year 3. *Counselor Education and Supervision, 45,* 162-179.

Milsom, A., & Moran, K. (2015). From school counselor to counselor educator: A phenomenological study. *Counselor Education and Supervision, 54,* 203-220.

Paglis, L. L., Green, S. G., Bauer, T. N. (2006). Does adviser mentoring add value? A longitudinal study of mentoring and doctoral student outcomes. *Research in Higher Education, 47,* 451-476.

Pasupathy, R., & Siwatu, K. O. (2014). An investigation of research self-efficacy beliefs and research productivity among faculty members at an emerging research university in the USA. *Higher Education Research & Development, 33,* 728-741.

Phillips, J. C., & Russell, R. K. (1994). Research self-efficacy, the research training environment, and research productivity among graduate students in counseling psychology. *The Counseling Psychologist, 22,* 628-641.

Peterson, R. L., Peterson, D. R., Abrams, J. C., Stricker, G., & Ducheny, K. (2010). The National Council of Schools and Programs of Professional Psychology Educational model – 2009. In M. B. Kenkel & R. L. Peterson (Eds.), *Competency-based education for professional psychology* (pp. 13-42). Washington, DC: American Psychological Association.

Potter, H., Higgins, G. E., & Gabbidon, S. L. (2011). The influence of gender, race/ethnicity, and faculty perceptions on scholarly productivity in criminology/criminal justice. *Journal of Criminal Justice Education, 22,* 84-101.

Richardon, V. (2006). Stewards of a field, stewards of an enterprise: The doctorate in education. In C. M. Golde, G. E. Walker, & Associates, *Envisioning the future of doctoral education: Preparing stewards of the disciplines* (Carnegie Essays on the Doctorate) (pp. 251-267). San Francisco, CA: Jossey-Bass.

Roussos ,L.A., Stout, W.F., & Marden, J.I. (1998). Using new proximity measures with hierarchical cluster analysis to detect multidimensionality. *Journal of Educational Measurement, 35* (1)*,* 1-30.

Walker, G. E., Golds, C. M., Jones, L., Bueschel, A. C., & Hutchings, P. (2008). *The formation of scholars: Rethinking doctoral education for the twenty-first century.* San Francisco, CA: Jossey-Bass.

Watson, J. C. (2017). Establishing evidence for internal structure using exploratory factor analysis. *Measurement and Evaluation in Counseling and Development, 50,* 232-238.

Wester, K. L., & Borders, L. D. (2014). Research competencies in counseling: A Delphi study. *Journal of Counseling & Development, 92,* 447-458.

Vera, M., Salanova, M., & Martin-del-Rio, B. (2011). Self-efficacy among university faculty: How to develop an adjusted scale. *Anales de Psicologia, 27,* 800-807.