

A Social Work Education Outcome Measure: The Evaluation Self-Efficacy Scale–II

SAGE Open
 July–September 2017: 1–9
 © The Author(s) 2017
 DOI: 10.1177/2158244017728323
journals.sagepub.com/home/sgo


Gary Holden¹, Kathleen Barker², Sofie Kuppens³,
 and Gary Rosenberg⁴

Abstract

The Evaluation Self-Efficacy (ESE) scale was designed as an outcome measure for evaluation research courses in social work. A revised version of the Social Cognitive Theory–based ESE (ESE-II) was developed and evaluated in the current study including both new and revised items. The ESE-II was evaluated in a final sample of 168 masters level students using a pretest–posttest design. Exploratory factor analysis revealed a single-factor structure underlying the 14 self-efficacy items at both assessment points. Cronbach’s alphas for the ESE-II were high at pre- and posttest. An argument underpinning content validity was developed and convergent validity was demonstrated. The ESE-II was also sensitive to change over time both at the item and scale level. The current study provides evidence supporting select psychometric properties of the ESE-II and the flexibility of self-efficacy as an outcome measure for social work education.

Keywords

assessment, evaluation, higher education, outcome measure, replication, self-efficacy, social work

Calls for greater use of evidence across many aspects of institutional practice increase by the day. To respond to these calls, education in the health related professions should develop a range of psychometrically sound outcome measures. Optimally these measures will be derived from empirically supported theories. There is substantial evidence supporting Albert Bandura’s Social Cognitive Theory and its component construct self-efficacy (e.g., Bandura, 1977, 1982, 1986, 1997, 2004, 2006a; Magklara, Burton, & Morrison, 2014; Shoji et al., 2016; Stajkovic & Luthans, 1998). Therefore, scales assessing self-efficacy regarding particular sets of professional behaviors may prove useful in meeting this need for outcome measures, especially in social work. This article describes the further development and evaluation of such a scale, namely, the Evaluation Self-Efficacy Scale–II (ESE-II).

- Education/training regarding: Administration in social work (York, 2008); skills of analysis in assessment (Platt, 2011); an evidence-supported treatment for community practice (Woody, Anderson, & D’Souza, 2015); child protection/child safeguarding (Carpenter, Patsios, Szilassy, & Hackett, 2011; Maxwell, Scourfield, Holland, Featherstone, & Lee, 2012; Scourfield, Smal, & Butler, 2014; Scourfield et al., 2012); child and family social work (Carpenter, Shardlow, Patsios, & Wood, 2013); undergraduate student writing skills (Woody et al., 2014); a core skills module for undergraduates and postgraduates (Tompsett, Henderson, Byrne, Mew, & Tomsett, 2017);
- Field instruction (Cuzzi, Holden, Chernack, Rutter, & Rosenberg, 1997; Cuzzi, Holden, Rutter, Rosenberg, & Chernack, 1996; Fortune, Lee, & Cavazos, 2005;

Self-Efficacy as an Educational Outcome in Social Work

The paucity of evidence-supported outcome measures in social work education has been apparent for some time (e.g., Bogo, Regehr, Hughes, Power, & Globerman, 2002). A growing body of research focusing on self-efficacy has begun to address this deficit. Subtopics such as the following have been examined:

¹New York University: Silver School of Social Work, New York, NY, USA

²Medgar Evers College of The City University of New York, Brooklyn, NY, USA

³Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

⁴Icahn School of Medicine at Mount Sinai, New York, NY, USA

Corresponding Author:

Gary Holden, New York University Silver School of Social Work, Room 409, 1 Washington Square North, New York, NY 10003, USA.
 Email: gary.holden@nyu.edu



- Holden, Cuzzi, Rutter, Rosenberg, & Chernack, 1996; Holden et al., 1997b; Holden, Cuzzi, Rutter, Chernack, & Rosenberg, 1997; Parker, 2005, 2006); including “shadowing” (Parker, Hughes & Rutter, 2006-2007);
- Social work research and evaluation (Holden, Barker, Meenaghan, & Rosenberg, 1999; Holden, Barker, Rosenberg, & Onghena, 2007, 2008; Macke & Tapp, 2012; Montcalm, 1999; Quinney & Parker, 2010; Unrau & Beck, 2004; Unrau & Grinnell, 2005); and
 - Social work educational outcomes in general (Ahn, Boykin, Hebert, & Kullkin, 2012; Holden, Anastas, & Meenaghan, 2003, 2005; Holden, Barker, Kuppens, & Rosenberg, 2017; Holden, Meenaghan, Anastas, & Metrey, 2002; Meyer-Adams, Dorsey, Potts, Resales, & Koob, 2011; Rawlings, 2012; Rishel & Majewski, 2009; Williams, King, & Koob, 2002).

This body of research suggests the versatility of self-efficacy as a construct (cf. Collins, 2015). It also supports our contention that the construct of self-efficacy is a strong basis for a wide-ranging, flexible system within which to assess outcomes in social work education (and likely other fields).

Replication

The importance of replication has been emphasized for many years (e.g., Ahlgren, 1969; Bornstein, 1990; Rosenthal, 1966, 1990; Smith, 1970). Yet, the prevalence of replication studies in social science journals in general and social work journals in particular remains low. The logical result is that calls for greater emphasis on, and more acceptance of, replication continue to be advanced across fields (e.g., Lindsay, 2015; Makel & Plucker, 2014; Warne, 2014).

Where does the current study fall in the range of potential replications? Fabrigar and Wegener (2015) noted,

[p]erhaps the categorization that has been most prevalent is the distinction between *exact* (often referred to as *direct*) and *conceptual* replication. The term *exact replication* typically refers to a study aimed at repeating as closely as possible the procedures used in a prior study. *Conceptual replication* is generally aimed at reproducing an effect from a prior study using different operationalizations of the independent (predictor) and/or dependent (outcome) variable(s) than in the prior study. (p. 68)

The current study (compared with the original ESE studies) assessed the same construct, in the same setting, with the same course instructor, over the same pre–post time frame, using the same research design and same scale format but with revised scale contents and a different sample.

Study Purposes

The current study serves two purposes (cf. Kaplan, Brownstein, & Graham-Day, 2017). First, it extends the development of an outcome measure designed for

accreditation purposes. Second, it seeks to provide faculty with a tool for curricular development activities beyond accreditation. To our knowledge, no other theoretically derived outcome measures are available that assess the outcomes of a course in the evaluation of social work practice. These ESE scales were designed to assess social work students’ self-efficacy regarding their ability to carry out evaluations of social work practice (e.g., quantitative, qualitative or mixed methods evaluations of social work practice). At the time the original ESE was created, the American Evaluation Association had not yet endorsed a group of competencies that could be used for item creation (Galport & Azzam, 2017). Because the objectives of any professional field or faculty within that field are subject to change, educational outcome measures must be able to efficiently adapt to those changes and not remain arrested over time. Given that the objectives of evaluation courses have evolved, the need to revise the original ESE to more effectively meet outcome assessment criteria became apparent.

This report details the results of a study of this revised ESE scale (ESE-II). Similar to earlier research on the ESE, the current study used a single-group pretest–posttest design, but recruited a different sample of students and tested the self-efficacy scale with revised and new items. Although the ESE was revised to accommodate course changes, we expected that the ESE-II would perform very similarly to the ESE upon which it is based. More specifically, given our prior experience with self-efficacy scales in general and social work self-efficacy scales designed to assess social work educational outcomes in particular, we sought to address the following research questions and predictions.

1. What is the underlying factor structure of the 14 self-efficacy items? As the factor structure of the ESE was not previously examined, we explored the underlying structure.
2. What is the reliability of the ESE-II? Similar to the original ESE studies (range $\alpha = .94-.96$), we predicted we would find high Cronbach’s alphas at both pretest and posttest for the ESE-II.
3. Validity of the ESE-II: (a) Do the ESE-II items cover the domain of evaluation research in social work, thus supporting content validity? (b) Is there evidence of convergent validity regarding the ESE-II? We sought to answer this question in two ways. First, the conceptual connection between empowerment and self-efficacy (including the use of self-efficacy as an indicator of empowerment) has a long tradition of discussion in social work (e.g., Evans, 1992; Gutiérrez, 1991; Gutiérrez & Lewis, 1999; Gutiérrez, Parsons, & Cox, 1998; Morton & Montgomery, 2012) as well as other fields (e.g., Crondahl & Karlsson, 2016). Based on our prior work with the ESE and the Social Worker Empowerment (SWE) scale (Frans, 1993), we predicted that the ESE-II

would have a small to moderate positive association with the SWE (comparable with the correlations found between the ESE and the SWE in the two prior studies). Second, we included three test self-efficacy items derived from the Council on Social Work Education's (CSWE; 2008) Educational Policy and Accreditation Standards (EPAS) for social work programs in the United States. Although we did not have prior data on the associations between these three items and the ESE, we assumed the correlation between the mean of these three items and the ESE-II total scale score would be positive and substantial.

4. In that the ESE was sensitive to change in the two prior studies, we predicted that the ESE-II would perform as well for both individual items and the total scale score.

Method

Study Design and Participants

This study used a single-group (across multiple cohorts), pretest–posttest design using a convenience sample acquired from 15 sections of a compulsory graduate course on evaluation. The sections were taught between the fall 2011 and summer 2015 semesters by the lead author in a large, urban, social work program in the Northeastern United States.

From a group of 210 participants whose anonymous scores were obtained at the pretest or posttest, a total of 168 complete usable sets of both pre- and posttest responses were available for psychometric analyses (i.e., no data were imputed; each respondent completed both pretest and posttest). At both pretest and posttest, students were presented with the option of participating, the procedure was explained and then the lead author left the room, while students did or did not complete the survey to maintain anonymity. No financial or other inducements for participation were offered. This study was approved as exempt by the University Institutional Review Board.

As these were *anonymous* surveys, no demographic data are available for the students in this study (see below). In terms of the larger population in this school from fall 2011 to summer 2015 (on the particular campus from which this convenience sample was obtained), the following demographic data are available. Eighty-five percent of the students were female with a mean age of 28.1 ($SD = 8.0$). International students comprised 8.5% of this student body. Major racial/ethnic categories were Asian/Pacific Islander (11.5%), Black (12.3%), Hispanic (12.9%), and White (56.9%).

Measures

SWE scale. Frans's (1993) 34-item, self-report SWE scale was used in prior studies of the ESE. In previous studies, Cronbach's alphas for the SWE ranged from .74 (Van

Voorhis & Hostetter, 2006) to .87 (Holden et al., 2007, 2008). In the current study, the Cronbach's alpha for the SWE was moderately high ($\alpha = .88$, $n = 196$).

ESE-II. The ESE and the ESE-II are both designed to assess social work students' self-efficacy regarding their ability to conduct evaluations of social work practice. They were developed following self-efficacy scale development guidelines (Bandura, 2006b). ESE instructions to participants were slightly revised for the ESE-II to enhance orientation of participants and their understanding of the scale instructions. In addition, items were modified to better reflect the complexity of the target behaviors. Next, three new items (Items 2, 4, and 9 in Table 1) were added. As can be seen in Table 1, the resulting 14-item scale has an 11-point response format (0 = *cannot do at all*; 50 = *moderately certain can do*; 100 = *certain can do*). The ESE-II total scale score is the mean of these 14 items.

Self-efficacy test items. Three self-efficacy test items were derived from the EPAS of CSWE (2008) and used in our *Self-Efficacy Regarding Social Work Competencies Scale* (Holden et al., in press). The items pertained to (a) using practice experience to inform scientific inquiry, (b) using research evidence to inform practice, and (c) social workers critically analyzing, monitoring, and evaluating interventions.

Procedure

Participants were assessed in the second and 13th session of a 14-session course. Administration was conducted by the instructor who explained that the instrument was voluntary and assured participants that he would not be in the room to monitor participation. He provided the directions for completing the instrument, answered any questions and then left the room. No demographic variables were assessed as this was an anonymous administration intended to further reduce socially desirable responding (Paulhus, 1991). We developed our approach for subject generated identification codes (i.e., *Hogben numbers*) before we discovered that similar ideas already existed in the literature (Honig, 1995; Yurek, Vasey, & Havens, 2008). This method allows matching of pre- and post-scores for individuals. This approach and additional details regarding the overall methodology used in these studies and justification for it are available from the first author. After each administration, data entry was carried out by a research assistant with a 10% check for errors by another research assistant.

Data-Analysis Strategy

The 14 items of the ESE-II were subjected to exploratory factor analysis (EFA) with principal axis factoring separately for pretest ($n = 207$) and posttest ($n = 169$) data to explore the factor structure of the scale. As an index of reliability, Cronbach's

Table 1. Item-Level Means and Standard Deviations for Pre- and Posttests for ESE-II (N = 168).

ESE-II item stem: How confident are you that you can . . .	Pretest	Posttest	Pre–post change	Cohen's
	M (SD) ^a	M (SD) ^b	M (95% CI) ^c	d ^e
1. Create an effective search strategy and conduct a thorough search of electronic library based databases, and other Internet resources, to obtain the scholarly literature necessary to design your evaluation of social work practice?	67.8^d (22.3)	81.2 (14.7)	12.9 (9.7-16.0)	.64
2. Critically examine and accurately describe tables and graphs found in the evaluation literature?	54.0 (22.4)	81.5 (13.6)	26.9 (23.6-30.1)	1.34
3. Critically review a particular area of social science theory and research (including the evaluation literature), and write a balanced and comprehensive literature review?	55.9 (22.2)	75.4 (17.6)	19.3 (16.0-22.6)	.91
4. Create a program logic model that clearly communicates the key factors and processes of social work practice?	35.0 (23.6)	71.1 (16.8)	36.6 (32.6-40.6)	1.41
5. Create and carry out a sampling strategy for your evaluation of social work practice?	45.6 (23.4)	77.3 (14.4)	32.9 (29.4-36.4)	1.51
6. Create and carry out a measurement approach for your evaluation of social work practice?	44.5 (23.9)	74.7 (14.2)	30.5 (26.9-34.0)	1.40
7. Create and carry out a descriptive data analysis plan (including data entry and management) for your evaluation of social work practice?	38.3 (25.8)	68.7 (17.1)	30.5 (26.9-34.2)	1.34
8. Create and carry out an inferential data analysis plan for your evaluation of social work practice?	30.3 (25.6)	65.4 (18.3)	36.0 (32.4-39.7)	1.57
9. Create and carry out a qualitative or mixed methods evaluation of social work practice?	42.2 (25.4)	75.9 (16.5)	33.9 (30.3-37.6)	1.50
10. Create and carry out a study regarding the implementation or processes of a social work program (e.g., needs assessment, formative evaluation, program description, program monitoring, quality assurance)?	45.6 (25.5)	76.2 (15.9)	30.7 (27.2-34.2)	1.46
11. Create and carry out a single system design to evaluate the outcomes of social work practice?	43.4 (25.4)	79.5 (15.2)	36.2 (32.5-39.9)	1.59
12. Create and carry out a group research design to evaluate the outcomes of social work practice?	42.0 (24.7)	75.4 (16.8)	33.9 (30.5-37.4)	1.57
13. Create and carry out an evaluation of social work practice that incorporates social work values and ethics (e.g., protects the participants in the evaluation)?	53.7 (24.6)	84.9 (13.1)	31.5 (27.8-35.1)	1.45
14. Create and carry out an evaluation of social work practice, while resisting political pressures from stakeholders that could bias your results or your evaluation report?	50.2 (28.0)	77.0 (15.7)	28.2 (24.3-32.2)	1.18
15. Use practice experience to inform scientific inquiry?	56.2 (25.5)	81.2 (16.1)		
16. Use research evidence to inform practice?	65.1 (21.1)	85.9 (13.4)		
17. Critically analyze, monitor, and evaluate interventions?	59.2 (22.0)	83.3 (13.8)		

Note. Higher scores indicate higher levels of self-efficacy. ESE-II = Evaluation Self-Efficacy Scale–II; CI = confidence interval.

^aPretest ESE-II *n*'s ranged from 208-210 due to missing data.

^bPosttest ESE-II *n*'s ranged from 169-172 due to missing data.

^cPre–post change scores and Cohen's *d* *n*'s were for the 168 complete cases.

^dHighest and lowest means are bolded.

alpha was computed with the lowest acceptable value set at .70 (Nunnally & Bernstein, 1994). Pearson correlations were used to examine convergent validity with criteria for a small, medium and large correlation set at .10, .30, and .50, respectively (Cohen, 1988). Pre–post change scores and 95% confidence interval (CI) were computed at ESE-II item and scale level in addition to standardized mean differences (*d*) for dependent groups. Cohen's *d* values of .30, .50, and .80 were considered to reflect a small, medium, and large difference, respectively (Cohen, 1988). Analyses were conducted using SPSS 24.0.

Results

Descriptive Statistics

All available data were used to compute descriptive statistics (*N*s ranged from 169 to 210). Following Cronbach's (1963) suggested focus on individual items as well as total scale scores, the specific item scores for students' evaluation self-efficacy on the ESE-II are presented in Table 1 (note that the sample sizes reflect all available respondents with usable data for each item).

The pretest item means ranged from 30.3 (i.e., *create and carry out an inferential data analysis plan for your evaluation of social work practice*) to 67.8 (i.e., *create an effective search strategy and conduct a thorough search of electronic library based databases, and other Internet resources, to obtain the scholarly literature necessary to design your evaluation of social work practice*).

The posttest item means ranged from 65.4 (i.e., *create and carry out an inferential data analysis plan for your evaluation of social work practice*) to 84.9 (i.e., *create and carry out an evaluation of social work practice that incorporates social work values and ethics (e.g., protects the participants in the evaluation)*). The pre–post change scores on the ESE-II and their 95% confidence intervals are also displayed in Table 1. The three self-efficacy test items ranged from 56.2 ($SD = 25.5$) to 65.1 ($SD = 21.1$) on the pretest and from 81.2 ($SD = 21.1$) to 85.9 ($SD = 13.4$) on the posttest.

Factor Structure

Prior to performing EFA with principal axis factoring, the suitability of the data was examined. The Kaiser-Meyer-Olkin (KMO) value of .93 ($>.60$) and a significant ($p < .001$) Bartlett's Test of Sphericity supported the factorability of the correlation matrix at both assessment points. Analyses pointed toward a single-factor solution at both assessment points as indicated by a clear break after the first factor in the scree plots. The single-factor solution explained 59.8% and 58.8% of the variance in the pre and posttest, respectively. All variables loaded substantially on the single factor with loadings ranging from .46 to .88 for the pretest and from .60 to .87 for the posttest. In both cases, Item 1 (*create an effective search strategy and conduct a thorough search of electronic library based databases, and other Internet resources, to obtain the scholarly literature necessary to design your evaluation of social work practice*) loaded least pronounced on the factor, while Item 12 (*create and carry out a group research design to evaluate the outcomes of social work practice*) yielded the strongest factor loadings.

Reliability

Analysis based on complete cases ($n = 168$) revealed high Cronbach's alphas for the ESE-II at pretest ($\alpha = .95$) and at posttest ($\alpha = .95$).

Validity

Content validity. Our argument in support of the content validity of the ESE-II is straightforward. Both the original ESE and the ESE-II use items reflecting the knowledge and skills one aims to improve in a course on evaluation in social work. For both scales, the items were directly matched to the objectives of the course in which data were obtained.

Construct validity. In line with our predictions, we found a positive correlation between the ESE-II and the SWE that was medium in magnitude, $r = .33$, according to Cohen's (1988) guidelines for interpreting effect sizes. To compare the ESE-II correlation with the SWE in the current study to the correlations between the ESE and SWE obtained in two prior studies ($r = .18$ and $r = .25$), we applied the estimation approach using Cumming's (2012) ESCI software (see also Association for Psychological Science, 2017; Cumming, 2015; Cumming & Calin-Jageman, 2017). The differences between the correlations were .15, 95% CI = $[-0.09, 0.4]$, and .08, 95% CI = $[-0.14, 0.30]$ with 95% confidence intervals including 0, indicating that the differences were not meaningful.

The positive correlation between the ESE-II and the mean of the three test self-efficacy items derived from the CSWE's (2008) EPAS was $r = .75$, 95% CI = $[0.68, 0.81]$, clearly exceeding Cohen's (1988) threshold of .50 reflecting a large effect size.

Sensitivity to Change

Pre–post change scores for individual items ranged from 12.9 to 36.6 (see Table 1). None of the 95% CIs for the individual pre–post change scores contained the value zero, suggesting that all of the particular aspects of student's self-efficacy over the course of the semester were strengthened. At item-level the pretest–posttest differences ranged from $d = .64$ to $d = 1.59$.

As presented in Table 2, the ESE-II total scale mean scores at pretest ranged from 5.7 to 99.3, in contrast to a minimum and maximum of 35.7 and 99.3 at posttest. Across students the pretest total scale mean was 45.9, 95% CI = $[42.9, 48.9]$, and the posttest mean was 75.9, 95% CI = $[74.0, 77.8]$. The difference between these total scale means was 30.0, 95% CI = $[27.4, 32.6]$ suggesting strengthening of students' self-efficacy over the course of the semester as the value zero was not included in the CI. In terms of Cohen's d (1988), the pretest–posttest difference was very large in magnitude ($d = 1.91$).

Discussion

The current study aimed at evaluating the psychometric properties of a revised tool to assess students' self-efficacy outcomes. Revision of the theoretically derived and empirically supported original ESE included rewording of directions and existing items. In addition, three new items were developed to better reflect the objectives of a revised course. Overall, findings supported the psychometric properties of the revised ESE-II evaluated in this study. EFA revealed a single-factor structure underlying the 14 self-efficacy items at the pre- and posttest assessment. The ESE-II scale also demonstrated excellent reliability at both assessment points. The argument made in the two prior studies—that the

Table 2. ESE-II Total Scale Scores ($n = 168$).

	Pretest	Posttest
<i>M</i>	45.9	75.9
<i>SD</i>	19.5	12.4
95% CI	[42.9, 48.9]	[74.0, 77.8]
Minimum-Maximum	5.7-99.3	35.7-99.3
Mean pre–post difference	30.0, 95% CI = [27.43, 32.57]	
Cohen's <i>d</i>	1.91	

Note. Higher scores indicate higher levels of self-efficacy. ESE-II = Evaluation Self-Efficacy Scale–II; CI = confidence interval.

original ESE could be seen as having preliminary evidence of content validity because the items were drawn directly from the objectives of a typical evaluation course—continues to apply in the current study. Preliminary evidence supporting the construct validity of the ESE-II was also observed. Finally, the ESE-II was sensitive to change in students' self-efficacy over the course of the semester, both at the item and scale level.

We frame the implications of these findings in the context noted by the Open Science Collaboration: “[i]nnovation points out paths that are possible; replication points out paths that are likely; progress relies on both. Replication can increase certainty when findings are reproduced and promote innovation when they are not” (2015, p. 943). Encouraging results across three studies increases certainty regarding the utility of the ESE and ESE-II. In addition, it continues to appear that Cronbach's (1963) suggested focus on individual items as well as total scale scores is wise. The substantial point difference on the lowest and highest rating at pretest leads to a number of questions related to curriculum that would be more obscure if the focus was on total scale scores alone. Modifying the ESE to produce the ESE-II to accommodate course changes, emphasizes the flexibility of the construct and supports our assertion that self-efficacy may form the basis of a larger outcome assessment system.

Although these results are consistent with prior research, some limitations should be noted. The sample of students was nonrandom and drawn from a single school of social work during a limited time period in the context of a single course. Although the current study replicates and extends our prior work by revising an existing scale to better assess the outcomes of a modified evaluation course, the psychometric properties of the ESE-II should be evaluated in additional situations. For instance, Deck, Platt, and Conner's (2015) recent assessment of a service learning research course demonstrated that the ESE performed similarly to our initial ESE studies.

The representativeness of the findings may have been reduced by attrition (20% reduction of useable scores). There are a variety of reasons for attrition possible here (e.g., refusals to participate, late for class or absent at the point of administration, leave of absence, shift into a different program track or to a different campus, dropped out, confused

by the identification system, lack of incentives for participation, no threat of being identified as a dropout, etc.). In comparison, our last five self-efficacy scale development studies (covering different pre–post period lengths) had attrition rates ranging from 21% to 45%. Thus, while the 20% attrition might have had an impact in this study, it falls at the lower end of attrition rates of our prior studies.

The use of self-efficacy in social work educational outcomes studies has been questioned on occasion (e.g., Drisko, 2014). As no one has challenged our earlier responses to these criticisms (Holden et al., 2017; Holden et al., 2007), we can only reiterate two key points. First, it appears some of those concerns flow from a basic misunderstanding of the construct. “Self-efficacy refers to beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments” (Bandura, 1997, p. 3). This definition is not equivalent to “a belief in one's ability to learn” (Drisko, 2014, p. 421). Neither is self-efficacy equivalent to self-esteem, self-worth, locus of control, or competence. Second, although Bandura has provided a guide for measuring self-efficacy (Bandura, 2006b), a number of recent studies in social work do not cite and/or optimally utilize this instruction (e.g., Jacobson, Osteen, Sharpe, & Pastoor, 2012; Oxhandler & Parrish, 2016; Parrish, Oxhandler, Duron, Swank, & Bordnick, 2016; Tompsett, Henderson, Byrne, Mew, & Tompsett, 2017). Adherence to conceptual and operational definitions is key when researchers utilize the construct. Moreover, failing to find a relationship between self-efficacy and outcomes with no supporting psychometric evidence *does not*, in our view, comprise a convincing argument against the use of self-efficacy.

In conclusion, it is logical to use theoretically derived, empirically supported constructs that have performed well across investigations. The current study provides additional evidence supporting the use of self-efficacy measures in the conduct of social work educational outcome assessments.

Authors' Note

The scale developed in this study is available at no charge from the first author.

Ethical Approval

New York University: Institutional Review Board approved as exempt.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Ahlgren, A. (1969). A modest proposal for encouraging replication. *American Psychologist*, *24*, 471. doi:10.1037/h0037798
- Ahn, B., Boykin, L., Hebert, C., & Kulkin, H. (2012). Exploring baccalaureate social work students' self-efficacy: Did it change over time? *Journal of Teaching in Social Work*, *32*, 487-501. doi:10.1080/08841233.2012.724046
- Association for Psychological Science (2017). *Advances in methods and practices in psychological science submission guidelines*. Retrieved from http://www.psychologicalscience.org/publications/psychological_science/ps-submissions#STAT
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review*, *84*, 191-215.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*, 122-147.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W. H. Freeman.
- Bandura, A. (2004). Swimming against the mainstream: The early years from chilly tributary to transformative mainstream. *Behaviour Research and Therapy*, *42*, 613-630.
- Bandura, A. (2006a). Toward a psychology of human agency. *Perspectives on Psychological Science*, *1*, 164-180.
- Bandura, A. (2006b). Guide for constructing self-efficacy scales. In F. Pajares & T. C. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Charlotte, NC: Information Age.
- Bogo, M., Regehr, C., Hughes, J., Power, R., & Globerman, J. (2002). Evaluating a measure of student field performance in direct service: Testing reliability and validity of explicit criteria. *Journal of Social Work Education*, *38*, 385-401.
- Bornstein, R. F. (1990). Publication politics, experimenter bias, and the replication process in social science research. *Journal of Social Behavior and Personality*, *5*, 71-81.
- Carpenter, J., Patsios, D., Szilassy, E., & Hackett, S. (2011). Outcomes of short course interprofessional education in parental mental illness and child protection: Self-efficacy, attitudes and knowledge. *Social Work Education: The International Journal*, *30*, 195-206. doi:10.1080/02615479.2011.540394
- Carpenter, J., Shardlow, S. M., Patsios, D., & Wood, M. (2013). Developing the confidence and competence of newly qualified child and family social workers in England: Outcomes of a national programme. *British Journal of Social Work*, *45*, 153-176. doi:10.1093/bjsw/bct106
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences, second edition*. Hillsdale, NJ: Lawrence Erlbaum.
- Collins, S. (2015). Alternative psychological approaches for social workers and social work students dealing with stress in the UK: Sense of coherence, challenge appraisals, self-efficacy and sense of control. *The British Journal of Social Work*, *45*(1), 69-85. doi:10.1093/bjsw/bct103
- Council on Social Work Education. (2008). *2008 Educational Policy and Accreditation Standards*. Retrieved from <https://www.cswe.org/Accreditation/Standards-and-Policies/2008-EPAS>
- Cronbach, L. (1963). Course improvement through evaluation. *Teachers College Record*, *64*, 672-683.
- Cron Dahl, K., & Karlsson, L. E. (2016). The nexus between health literacy and empowerment: A scoping review. *SAGE Open*, *6*(2), 1-7. doi:10.1177/2158244016646410
- Cumming, G. (2012). *Exploratory software for confidence intervals, ESCI*. Retrieved from <http://www.latrobe.edu.au/psychology/research/research-areas/cognitive-and-developmental-psychology/esci>
- Cumming, G. (2015). The new statistics: Why and how. *Psychological Science*, *25*, 7-29. doi:10.1177/0956797613504966
- Cumming, G., & Calin-Jageman, R. (2017). *Introduction to the new statistics: Estimation, open science and beyond*. New York, NY: Routledge.
- Cuzzi, L. C., Holden, G., Chernack, P., Rutter, S., & Rosenberg, G. (1997). Evaluating social work field instruction: Rotations versus year-long placements. *Research on Social Work Practice*, *7*, 402-414. doi:10.1177/104973159700700308
- Cuzzi, L. C., Holden, G., Rutter, S., Rosenberg, G., & Chernack, P. (1996). A pilot study of fieldwork rotations vs. year long placements for social work students in a public hospital. *Social Work in Health Care*, *24*, 73-91.
- Deck, S. M., Platt, P. A., & Conner, L. (2015). Engaged teaching-learning: Outcome evaluation for social work students in a graduate-level service learning research course. *Advances in Social Work*, *16*, 233-248. doi:10.18060/18302
- Drisko, J. (2014). Competencies and their assessment. *Journal of Social Work Education*, *50*, 414-426. doi:10.1080/10437797.2014.917927
- Evans, E. N. (1992). Liberation theology, empowerment theory and social work practice with the oppressed. *International Social Work*, *35*, 135-147.
- Fabrigar, L. R., & Wegener, D. T. (2015). Conceptualizing and evaluating the replication of research results. *Journal of Experimental Social Psychology*, *66*, 68-80. doi:10.1016/j.jesp.2015.07.009
- Fortune, A. E., Lee, M., & Cavazos, A. (2005). Achievement motivation and outcome in social work field education. *Journal of Social Work Education*, *41*, 115-129. doi:10.5175/JSWE.2005.200300318
- Frans, D. J. (1993). A scale for measuring social worker empowerment. *Research on Social Work Practice*, *3*, 312-328.
- Galport, N., & Azzam, T. (2017). Evaluator training needs: A gap analysis. *American Journal of Evaluation*, *38*, 80-100. doi:10.1177/1098214016643183
- Gutiérrez, L. (1991). Empowering women of color: A feminist model. In M. Bricker-Jenkins, N. R. Hooyman, & N. Gottlieb (Eds.), *Feminist social work practice in clinical settings* (pp. 199-214). Newbury Park, CA: SAGE.
- Gutiérrez, L., & Lewis, E. A. (1999). Empowerment: A model for practice. In L. M. Gutiérrez & E. A. Lewis (Eds.), *Empowering women of color* (pp. 3-23). New York, NY: Columbia University Press.
- Gutiérrez, L., Parsons, R. J., & Cox, E. O. (1998). A model for empowerment practice. In L. M. Gutiérrez, R. J. Parsons, & E. O. Cox (Eds.), *Empowerment in social work practice: A sourcebook* (pp. 3-23). Pacific Grove, CA: Brooks/Cole.
- Holden, G., Anastas, J., & Meenaghan, T. (2003). Determining attainment of the EPAS foundation program objectives: Evidence for the use of self-efficacy as an outcome. *Journal of Social Work Education*, *39*, 425-440.
- Holden, G., Anastas, J., & Meenaghan, T. (2005). EPAS objectives and foundation practice self-efficacy: A replication. *Journal of Social Work Education*, *41*, 559-570.

- Holden, G., Barker, K., Kuppens, S., & Rosenberg, G. (2017). Self efficacy regarding social work competencies. *Research on Social Work Practice*, 27, 594-606. doi:10.1177/10497315155586466
- Holden, G., Barker, K., Meenaghan, T., & Rosenberg, G. (1999). Research self-efficacy: A new possibility for educational outcomes assessment. *Journal of Social Work Education*, 35, 463-476.
- Holden, G., Barker, K., Rosenberg, G., & Onghena, P. (2007). Assessing progress towards accreditation related objectives: Evidence regarding the use of self-efficacy as an outcome in the advanced concentration research curriculum. *Research on Social Work Practice*, 17, 456-465.
- Holden, G., Barker, K., Rosenberg, G., & Onghena, P. (2008). The Evaluation Self-Efficacy Scale: A replication. *Research on Social Work Practice*, 18, 42-46.
- Holden, G., Cuzzi, L. C., Rutter, S., Chernack, P., & Rosenberg, G. (1997). The Hospital Social Work Self-Efficacy Scale: A replication. *Research on Social Work Practice*, 7, 490-499.
- Holden, G., Cuzzi, L. C., Rutter, S., Chernack, P., Spitzer, W., & Rosenberg, G. (1997b). The Hospital Social Work Self-Efficacy Scale: A partial replication and extension. *Health & Social Work*, 22, 256-263.
- Holden, G., Cuzzi, L. C., Rutter, S., Rosenberg, G., & Chernack, P. (1996). The Hospital Social Work Self-Efficacy Scale: Initial development. *Research on Social Work Practice*, 6, 353-365.
- Holden, G., Meenaghan, T., Anastas, J., & Metrey, G. (2002). Outcomes of social work education: The case for social work self-efficacy. *Journal of Social Work Education*, 38, 115-133.
- Honig, F. (1995). When you can't ask their names: Linking anonymous respondents with the Hogben number. *Australian Journal of Public Health*, 19, 94-96. doi:10.1111/j.1753-6405.1995.tb00305.x
- Jacobson, J. M., Osteen, P. J., Sharpe, T. L., & Pastoor, J. B. (2012). Randomized trial of suicide gatekeeper training for social work students. *Research on Social Work Practice*, 22, 270-281.
- Kaplan, C. S., Brownstein, E. M., & Graham-Day, K. J. (2017). One for all and all for one: Multi-university collaboration to meet accreditation requirements. *SAGE Open*, 7(1), 1-13. doi:10.1177/2158244016687610
- Lindsay, D. S. (2015). Replication in psychological science. *Psychological Science*, 26, 1827-1832. doi:10.1177/0956797615616374
- Macke, C., & Tapp, K. (2012). Teaching research to MSW students: Effectiveness of the team-based learning pedagogy. *Journal of Teaching in Social Work*, 32, 148-160. doi:10.1080/08841233.2012.668161
- Magklara, E., Burton, C. R., & Morrison, V. (2014). Does self-efficacy influence recovery and well-being in osteoarthritis patients undergoing joint replacement? A systematic review. *Clinical Rehabilitation*, 28, 835-846. doi:10.1177/0269215514527843
- Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher*, 43, 304-316. doi:10.3102/0013189X14545513
- Maxwell, N., Scourfield, J., Holland, S., Featherstone, B., & Lee, J. (2012). The benefits and challenges of training child protection social workers in father engagement. *Child Abuse Review*, 21, 299-310. doi:10.1002/car.2218
- Meyer-Adams, N., Dorsey, C. J., Potts, M. K., Resales, A. M., & Koob, J. J. (2011). How to tackle the shift of educational assessment from learning outcomes to competencies: One program's transition. *Journal of Social Work Education*, 47, 489-507. doi:10.5175/JSWE.2011.201000017
- Montcalm, D. M. (1999). Applying Bandura's theory of self-efficacy to the teaching of research. *Journal of Teaching in Social Work*, 19, 93-107. doi:10.1300/J067v19n01_08
- Morton, M. H., & Montgomery, P. (2012). Youth empowerment programs for improving adolescents' self-efficacy and self-esteem. *Research on Social Practice*, 23, 22-33. doi:10.1177/1049731512459967
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349, 6251, 943. doi:10.1126/science.aac4716
- Oxhandler, H. K., & Parrish, D. E. (2016). The development and validation of the Religious/Spiritually Integrated Practice Assessment Scale. *Research on Social Work Practice*, 26, 295-307. doi:10.1177/1049731514550207
- Parker, J. (2005). Should you encourage students to assess themselves in practice learning? A guided self-efficacy approach to practice learning assessment. *Journal of Practice Teaching*, 6(3), 8-30.
- Parker, J. (2006). Developing perceptions of competence during practice learning. *The British Journal of Social Work*, 36, 1017-1036. doi:10.1093/bjsw/bch347
- Parker, J., Hughes, M., & Rutter, L. (2006-2007). "Shadowing" and its place in preparing students for practice learning. *The Journal of Practice Teaching & Learning*, 7, 51-70. doi:10.1921/19643
- Parrish, D. E., Oxhandler, H. K., Duron, J. F., Swank, P., & Bordnick, P. (2016). Feasibility of virtual reality environments for adolescent social anxiety disorder. *Research on Social Work Practice*, 26, 825-835. doi:10.1177/1049731514568897
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-60). New York, NY: Academic Press.
- Platt, D. (2011). Assessments of children and families: Learning and teaching the skills of analysis. *Social Work Education*, 30, 157-169. doi:10.1080/02615479.2011.540381
- Quinney, A., & Parker, J. (2010). Developing self efficacy in research skills: Becoming research-minded. In H. Burgess & J. Carpenter J. (Eds.), *The outcomes of social work education: Developing evaluation methods* (pp. 16-25). Southampton, UK: Higher Education Academy Subject Centre for Social Policy and Social Work.
- Rawlings, M. A. (2012). Assessing BSW student direct practice skill using standardized clients and self-efficacy theory. *Journal of Social Work Education*, 48, 553-576. doi:10.5175/JSWE.2012.201000070
- Rishel, C. W., & Majewski, V. (2009). Student gains in self-efficacy in an advanced MSW curriculum: A customized model for outcomes assessment. *Journal of Social Work Education*, 45, 365-383. doi:10.5175/JSWE.2009.200800101
- Rosenthal, R. (1966). *Experiment effects in behavioral research*. New York, NY: Appleton-Century-Crofts.
- Rosenthal, R. (1990). Replication in behavioral research. *Journal of Social Behavior and Personality*, 5, 1-30.
- Scourfield, J., Smail, P., & Butler, D. (2014). Improving the engagement of fathers in child safeguarding. *Child Abuse Review*, 24, 77-153. doi:10.1002/car.2333

- Scourfield, J., Tolman, R., Maxwell, N., Holland, S., Bullock, A., & Sloan, L. (2012). Results of a training course for social workers on engaging fathers in child protection. *Children & Youth Services Review, 34*, 1425-1432. doi:10.1016/j.childyouth.2012.03.022
- Shoji, K., Cieslak, R., Smoktunowicz, E., Rogala, A., Benight, C. C., & Luszczynska, A. (2016). Associations between job burnout and self-efficacy: A meta-analysis. *Anxiety, Stress, & Coping: An International Journal, 29*, 367-386. doi:10.1080/10615806.2015.1058369
- Smith, N. C. (1970). Replication studies: A neglected aspect of psychological research. *American Psychologist, 25*, 970-975. doi:10.1037/h0029774
- Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124*, 240-261.
- Tompsett, H., Henderson, K., Byrne, J. M., Mew, E. G., & Tomsett, C. (2017). Core-skills module for all students, both postgraduate (PG) and undergraduate (UG), on two social work programmes. *The British Journal of Social Work, 1*-22.
- Unrau, Y. A., & Beck, A. R. (2004). Increasing research self-efficacy among students in professional academic programs. *Innovative Higher Education, 28*, 187-204.
- Unrau, Y. A., & Grinnell, R. M. (2005). The impact of social work research courses on research self-efficacy for social work students. *Social Work Education, 24*, 639-651. doi:10.1080/0261547050018506
- Van Voorhis, R. M., & Hostetter, C. (2006). The impact of MSW education on social worker empowerment and commitment to client empowerment through social justice advocacy. *Journal of Social Work Education, 42*, 105-121.
- Warne, R. T. (2014). Two additional suggested reforms to encourage replication studies in educational research. *Educational Researcher, 43*, 465. doi:10.3102/0013189X14562294
- Williams, N. R., King, M., & Koob, J. J. (2002). Social work students go to camp: The effects of service learning on perceived self-efficacy. *Journal of Teaching in Social Work, 22*, 55-70. doi:10.1300/J067v22n03_05
- Woody, J. D., Anderson, D. K., & D'Souza, H. J. (2015). Dissemination of trauma-focused cognitive-behavioral therapy with community practitioners: Focus on self-efficacy. *Journal of Evidence-Informed Social Work, 12*, 289-301. doi:10.1080/5433714.2014.950128
- Woody, J. D., Zeleny, M. G., D'Souza, H. J., Harder, J., Reiser, J., & Szto, P. (2014). Student progress in a social work writing course: Self-efficacy, course objectives, and skills. *Journal of Social Work Education, 50*, 507-524. doi:10.1080/10437797.2014.91789
- York, R. O. (2008). Comparing three modes of instruction in a graduate social work program. *Journal of Social Work Education, 44*, 157-172.
- Yurek, L. A., Vasey, J., & Havens, D. S. (2008). The use of self-generated identification codes in longitudinal research. *Evaluation Review, 32*, 435-452.

Author Biographies

Gary Holden is a professor in the Silver School of Social Work at New York University. He teaches an evaluation course and his primary research area focuses on educational outcomes assessment. He developed and maintains the social work knowledge dissemination site Information for Practice (<http://ifp.nyu.edu/>).

Kathleen Barker is a professor of Psychology, Medgar Evers College of The City University of New York. She teaches courses on methodology, social psychology, and applied areas.

Sofie Kuppens is an assistant professor at the Erasmus School of Health Policy and Management (ESHPM). She teaches courses in research methods and quantitative analysis in the social and health sciences.

Gary Rosenberg is the Edith J. Baerwald professor at the Icahn School of Medicine and the editor in chief of the journal *Social Work In Health Care*.