

A Study of the Relationship between Educational Self-Efficacy and Educational Performance
among Social Work Students

Self-efficacy is the individual's perception of ability to successfully engage in certain tasks, such as interviewing a client using non-directive techniques, designing an evaluation study that appropriately incorporates certain research standards, or preparing a policy brief on a social policy issue. It refers to specific situations rather than self-confidence in general. It is a concept that has received growing attention in social work education as a measure of outcome, based on the assumption that self-efficacy is an indicator of competence (Holden, Barker, Rosenberg, & Onghena, 2007; Holden, Meenaghan, Anastas, & Metrey, 2002). The ability of a measure of self-efficacy to predict performance based on other measures of performance is a means of testing the assumption that self-efficacy can be a way of measuring competence. The research study reported here examines the relationship between evaluation research self-efficacy and evaluation research knowledge.

Self-efficacy is measured by self-perception, so its validity relies upon the accuracy of an individual's self-perception. This issue has been challenged by some research suggesting that self-efficacy scores are lower when measured retrospectively than the traditional format for the pretest measurement. In the retrospective format for the pretest, study subjects are asked at the posttest time to reflect upon where they were at the pretest time and to complete the self-efficacy instrument with this thought in mind. The result has been that pretest scores are lower when measured retrospectively than traditionally (Holden, Anastas, & Meenaghan, 2005).

One of the questions pursued in the present study is whether the retrospective pretest score is a valid measure of pretest conditions. In the study reported here, gain on self-efficacy using the retrospective pretest is compared with gain in self-efficacy using a traditional pretest as

a predictor of gain on performance using another measure of performance. For example, gain in research knowledge using a pretest and posttest exam will be correlated with gain in self-efficacy using both ways to measure self-efficacy. Is the correlation higher for one than the other?

Gain in knowledge in research for social work students is the measure of competence for the research reported here. Students in several sections of research courses were measured before and after a research course on knowledge and self-efficacy. The key question is whether higher knowledge is associated with higher self-efficacy.

In one set of studies undertaken in the [spring](#) of 2014, students in several sections of a second undergraduate research course and one section of a second graduate research course were given tools to measure knowledge, self-efficacy, gain on knowledge, and gain on self-efficacy. In a second set of studies undertaken in the [fall](#) of 2014, students in several sections of a [first](#) undergraduate research course and one section of a first graduate research course were measured on the same variables.

Among the research questions pursued were:

1. Are evaluation research self-efficacy pretest scores positively correlated with evaluation research pretest knowledge, when the self-efficacy score is measured in the traditional method for the pretest?
2. Are evaluation research self-efficacy pretest scores positively correlated with evaluation research pretest knowledge, when the self-efficacy score is measured in the retrospective method for the pretest?
3. Are evaluation research self-efficacy posttest scores positively correlated with evaluation research posttest knowledge?

Commented [U1]: Perhaps rephrase into a statement: This study sought to determine whether the correlation is higher for one than the other.

Commented [U2]: Is there a reason these are called first while the previously mentioned classes are called second? I might just not be familiar with the terminology.

4. Are evaluation research self-efficacy gain scores positively correlated with evaluation research knowledge gain scores, when the gain is measured using the traditional pretest method?
5. Are evaluation research self-efficacy gain scores positively correlated with evaluation research knowledge gain scores, when the gain is measured using the retrospective pretest method?

Self-efficacy as a Measure of Educational Outcome

Bandura (1977; 1997) defined self-efficacy as the belief that one is capable of acting in certain ways in order to produce a desired result. Higher self-efficacy may be indicative of motivation to act, specifically in ways that will achieve the desired outcome. Therefore, those with lower self-efficacy regarding a specific task may be less motivated to attempt the task, and expend less effort at the task, thereby ensuring less than optimal results. Self-efficacy has been measured in a variety of ways: as a global construct, as measured in the Self-Efficacy Scale (Sherer et. al, 1982); and in a variety of measures more specific to a particular task, such as health behaviors and athletic performance (Holden, 1991; Wise & Trunnell, 2001).

The measurement of self-efficacy in regard to a specific domain has proved more useful than measuring it broadly. Choi (2005) reports that research involving self-efficacy has been hampered by attempts to measure the construct in a global fashion, and notes that “According to Bandura (1997) and Pajares (1996) this lack of correspondence between the level of self-efficacy measured and the criterial performance in research studies is one reason that some researchers do not observe a significant effort of the self-efficacy construct on the dependent variable measure.” (Choi, 2005, p 197).

The relationship of self-efficacy and human performance has been well-researched across professions ~~related to such fields as such as~~ health, psychology, and ~~professions related to~~ athletic performance. A key word search of 'self-efficacy' and 'academic performance' in Pro Quest Academic elicited over 1,000 peer-reviewed research articles. As measurement of academic performance is central to this study, a search in databases related to research in education and applied psychology was conducted. Results tend to be mixed and varied.

One study found that students who entered college with higher confidence and higher expectations performed significantly better than less confident ~~ee~~ students (Chemers, Hu, and Garcia, (2001). Indeed, this relationship between self-efficacy and performance appears commonly in the literature. A study by Finney and Schraw (2003) found that self-efficacy was related to performance in a statistics course. A meta-analysis of research in educational settings found that self-efficacy was related to both academic performance and ~~to~~ persistence (Multon, Brown, & Lent, 1999).

Self-efficacy for social work practice and social work research has been the subject of a number of articles that reported the testing of various self-efficacy scales that targeted selected aspects of social work practice. One of these is The Hospital Social Work Self-efficacy Scale (Holden, Cuzzi, Spitzer, Rutter, Chernack, & Rosenberg, 1997) and one of these is The Evaluation Self-efficacy Scale (Holden, Barker, Rosenberg, & Onghena, 2007; 2008). Each of these scales asks respondents to indicate their level of confidence that they can perform certain tasks at a high level. The score for each of the tasks in a general category are summed for a score. The Evaluation Self-Efficacy Scale is the key one used in the present study.

There is scant literature on self-efficacy as it relates to performance in social work. With regard to social work student self-efficacy, satisfaction with field practicum, and self-rated abilities, Fortune, Lee, and Cavazos (2005) found that self-efficacy was related to satisfaction and perception of ability. However, findings were not considered with regard to a performance outcome.

Rawlings (2012) examined the relationship of self-efficacy and performance in an interview of a standardized client for both beginning and graduating BSW students. The findings indicated that direct practice skill and direct practice self-efficacy were not significantly correlated. Moreover, when other variables were controlled, it was found that direct practice self-efficacy and direct practice skill were negatively related at a significant level. Both direct practice skill and direct practice self-efficacy were higher for those ending their BSW education than for those just beginning. This would suggest that both self-efficacy and skill are improved with BSW education, but these two variables are not related in the way we would expect. These findings challenge the use of self-efficacy as a substitute for the measurement of human performance.

The question that is central to the present article is whether self-efficacy is a legitimate measure of human performance. If so, it can be employed as such a measure. Several articles have proposed self-efficacy as a measure of educational outcome for schools of social work and several schools of social work have adopted this practice. If self-efficacy and educational outcomes (measured more directly) are not related, however, we would have reason to question this practice.

Study Methods

One of the variables measured in this study was evaluation research self-efficacy. This variable was measured by the Evaluation Self-efficacy Scale developed and tested by Holden and others (Holden, Barker, Rosenberg, & Onghena, 2008). This scale has 11 items, each of which portrays an evaluation research competence. For each of the items, the respondent is asked “How confident are you that you can successfully [complete each task]?”. One of these items is “create a group research design to evaluate the outcomes of practice” and another is “design an inferential data analysis for your evaluation of some aspect of practice” while a third one is “review a particular area of social science research (including the evaluation literature), write a balanced and comprehensive literature review.” This scale was given at the beginning of the class experience and again at the end. A revised form of this scale was also employed at the end of the class whereby students were asked to reflect upon where they were at the beginning of the class experience. The instructions said, in part, “We want to know how confident you were, in your ability to perform specific social work tasks before you took this course.” At the top of the page that presented the 11 evaluation research competencies was the following: “Before this course experience, how confident were you that you could have successfully....” On each item, the respondent circles a number between 0 and 100, so the maximum possible score is 1100 and the minimum score is 0, with higher scores indicating a higher level of self-efficacy.

Another variable was evaluation research knowledge. This was measured by a test designed by one of the co-authors of this article who taught three of the five classes that were included in this study. The exam for the first study (SpringSpring 2014) was obviously too difficult for the study being undertaken. One of the problems with this exam was that it presented a research example and the exam called upon the student to go back to the example when reviewing each question. This problem, coupled with the fact that the exam was voluntary,

Formatted: Font: Not Bold

Commented [CM3]: I don't think I would underline this, since it's the first time underlining appears—maybe put in quotation marks if that would keep consistency.

caused a number of students not to complete it and others to hurry through it after a few minutes. One student was heard to say, “This exam is making my head hurt.” Consequently, the exam was revised for the second study.

The second exam had 20 questions. The competencies measured by the odd numbered items on this multiple choice exam were the following:

1. The ability to distinguish between statements of the purpose of an evaluation study that were and were not consistent with the spirit of scientific inquiry.
2. The ability to identify the threats to internal validity that were addressed by the one-group pretest posttest research design.
3. The ability to recognize that the chief advantage of a good analysis of target behavior in evaluative research is guidance on the selection of an intervention.
4. The ability to identify the best way to articulate the hypothesis for an evaluative research study.
5. The ability to determine if the hypothesis was supported in an example when given the data results.
6. The ability to identify the dependent variable in an evaluative research example.
7. The ability to distinguish between the situations that would better address a study question using qualitative research measurement rather than quantitative research measurement.
8. The basic issue that is addressed by the statistical analysis of data.
9. The ability to identify the issue of justice in research ethics as a major concern when given an example.
10. The ability to identify how many variables are in a descriptive analysis of data.

There were five evaluation research classes that participated in this study. All five of these classes were in the same university. One of the classes was a first research course in the social work graduate program. All of the other four classes were undergraduate research classes, two taught by the senior author of this article and the other two taught by the other authors of this article. The graduate course was taught by the senior author.

There were two iterations of this study, one conducted in the [Spring](#) Semester of 2014 and the other conducted in the [Fall](#) Semester of 2014. The second iteration was necessitated by the fact that the students in the first study failed to demonstrate a gain in knowledge, so that knowledge gain failed to be a good variable for testing the hypotheses of the study. This problem was remedied with the second study. The results of both will be reported.

The key study hypotheses tested were:

1. Posttest scores for evaluation research knowledge (exam) will be higher than pretest scores.
2. Posttest scores for evaluation self-efficacy will be higher than pretest scores.
3. Posttest scores for evaluation self-efficacy will be higher than retrospective pretest scores.
4. Pretest scores for knowledge will be positively correlated with pretest scores for evaluation self-efficacy when pretest scores for self-efficacy are measured in the traditional way.
5. Pretest scores for knowledge will be positively correlated with pretest scores for evaluation self-efficacy when self-efficacy is measured retrospectively.
6. Posttest scores for evaluation research knowledge will be positively correlated with posttest scores for evaluation self-efficacy.
7. Gain on knowledge will be positively correlated with gain on evaluation self-efficacy when gain on evaluation self-efficacy is measured by the tradition pretest.

8. Gain on knowledge will be positively correlated with gain on evaluation self-efficacy when self-efficacy gain is measured with the retrospective pretest method.

Results

The first part of this report of results will be for the study conducted in the [Spring](#) of 2014. The key questions focused [on](#) the relationship between self-efficacy and research knowledge. However, as detailed in the first result reported below, the failure of the students to achieve a gain on evaluation research knowledge proved to be a problem for the testing of most of the hypotheses. Here are the results in regard to the hypotheses stated above.

1. The average exam score for the pretest was 36.96 while the average exam score for the posttest was lower (34.58), so the students did not achieve a gain. This rendered the examination of gain scores as problematic for the examination of the study hypotheses.
2. The average posttest score for evaluation self-efficacy (844.75) was significantly lower than the average pretest score for evaluation self-efficacy (619.96) when pretest was measured in the traditional way ($t = 8.36$, $df = 36$; $p < .001$).
3. The average posttest score for evaluation self-efficacy (844.75) was significantly lower than the average pretest score for evaluation self-efficacy (431.31) when pretest was measured in the retrospective way ($t = 14.09$, $df = 36$; $p < .001$).
4. There was not a significant correlation between pretest exam scores and pretest self-efficacy scores when the pretest was taken in the traditional method ($r = -.02$; $n = 62$; $p = .85$). There was not a significant correlation between pretest exam scores and pretest self-efficacy scores when the pretest was taken in the retrospective method ($r = -.004$; $n = 36$; $p = .98$).

5. There was not a significant correlation between posttest exam scores and posttest self-efficacy scores ($r = -.20$; $n = 33$; $p = .25$).
6. There was not a significant correlation between exam score gain and self-efficacy gain when the pretest was measured in the traditional fashion ($r = .23$; $n = 32$; $p = .20$).
7. There was not a significant correlation between exam score gain and self-efficacy gain when the pretest was measured in the retrospective format ($r = .27$; $n = 32$; $p = .13$).

These results fail to show a relationship between knowledge and self-efficacy, nor do they show a relationship between gain in knowledge and gain on self-efficacy. However, the failure of these students to demonstrate a gain in knowledge renders this analysis problematic. Because the research course given in this study was the second research course, the exam was designed to be difficult in order to avoid having a pretest score that was too high.

There ~~was~~ were, however, lessons learned from this study. The fact that pretest scores were lower when taken retrospectively than traditionally is similar to the results of other studies. This finding suggests that students are more aware of what they did not know at the beginning of the experience after they have had the learning experience than they ~~did~~ were at the beginning of the learning experience.

The second study pursued the same hypotheses with a different instrument for measuring evaluative research knowledge. The results are as follows:

1. Students achieved a significant gain in evaluation research knowledge. The mean pretest score for the evaluation research exam was 29.5 while the posttest was 41.7. This difference was statistically significant ($t = 6.5$; $df = 86$; $p < .001$).
2. Students achieved a gain in evaluation research self-efficacy when the pretest was measured traditionally. The mean pretest (traditional pretest) score for evaluation self-

efficacy was 541.35 while the mean for the posttest was 747.65. This difference was significant ($t = 7.73$; $df = 80$; $p < .001$).

3. Students demonstrated a gain in evaluation research self-efficacy when the pretest was measured retrospectively. The mean retrospective pretest score was 473.44 while the mean posttest was 755.63 ($t = 11.5$; $df = 86$; $p < .001$). Because the retrospective pretest was lower than the traditional pretest, the amount of gain was greater when the retrospective pretest measure was employed.
4. There was not a significant correlation between pretest exam scores and pretest self-efficacy scores when the pretest was taken in the traditional method ($r = .088$; $n = 86$; $p = .418$).
5. There was not a significant correlation between pretest exam scores and pretest self-efficacy scores when the pretest was taken in the retrospective method ($r = .159$; $n = 92$; $p = .13$).
6. There was not a significant correlation between posttest exam scores and posttest self-efficacy scores ($r = .20$; $n = 83$; $p = .069$).
7. There was not a significant correlation between exam score gain and self-efficacy gain when the pretest was measured in the traditional fashion ($r = .03$; $n = 76$; $p = .78$).
8. There was not a significant correlation between exam score gain and self-efficacy gain when the pretest was measured in the retrospective format ($r = .05$; $n = 81$; $p = .65$).

There were five different classes of students included in this study. The question of whether there were differences between classes was undertaken by a one-way analysis of variance with class as the independent variable. The results are displayed in Table 1.

Insert Table 1 here.

One of the results to be expected was that the variables for the graduate class tended to be higher than those for the undergraduate classes ~~on these variables~~. A finding not expected was that the undergraduate class of students in the extension program tended to be lower than the other classes in regard to self-efficacy, but they were not lower than the other undergraduate classes in regard to research course performance (exam scores). In fact, this extension class was the lowest in self-efficacy for all undergraduate classes, but ~~they were~~ they tended to be slightly higher than the other undergraduate classes in regard to exam scores both pretest and posttest. This specific difference, however, was not subjected to statistical analysis, and, given the small differences, we should probably say we failed to find a difference. This finding would suggest there are variables other than performance that explain self-efficacy. Perhaps, for example, the extension students have lower self-esteem in regard to academic performance given the fact they tend to be community college transfer students.

Given the differences between the graduate class and the other classes, it was decided to undertake a testing of the eight study hypotheses for this class alone. The results were similar to the general findings of this study. There was not a significant correlation between pretest knowledge and pretest self-efficacy ($r = .36; n = 18; p = .14$), between posttest knowledge and posttest self-efficacy ($r = .18; n = 18; p = .48$), between knowledge gain and self-efficacy gain (~~$r = .11; n = 17; p = .69$~~) when the gain for self-efficacy was measured using the traditional method ($r = -.11; n = 17; p = .69$), or between knowledge gain and self-efficacy gain when gain was measured using the retrospective method ($r = .17; n = 18; p = .50$). There was one exception

Commented [CM4]: Moved for consistent placement—OK?

to this pattern. For this graduate class, there was a strong positive correlation between exam pretest scores and self-efficacy pretest scores when self-efficacy was measured with the retrospective method ($r = .71$; $n = 19$; $p = .001$). This is the one difference between the graduate class and the results for the entire study.

To complete the examination of potential differences among the classes, a set of dummy variables were composed with one variable for each class which compared that particular class with all the others in the study sample. For each of these variables the targeted class was coded as 1 while all others were coded as 0. This set of analyses did not challenge the general findings. What was found was that (a) graduate students were generally higher than others on exam scores and self-efficacy, and (b) the main campus undergraduate students taught by the senior author of this study scored higher on self-efficacy than the others. The latter fact was unusual in that this undergraduate class (taught by the senior author) did not perform better on the exam than the other class taught by the senior author.

Discussion

The basic research question for this study was as follows: Is evaluation research knowledge positively related to evaluation research self-efficacy? The findings of this study suggest that the answer is **No**. Knowledge was not found to be correlated with self-efficacy at either the pretest time or the posttest time, and knowledge gain was not found to be correlated with self-efficacy gain. These findings held true when self-efficacy was measured in the traditional format or in the retrospective format.

There are a few other findings of this study that are noteworthy. First, pretest self-efficacy scores were found to be higher than retrospective pretest self-efficacy scores. This finding held true for all five of the classes of students in this study. The explanation that makes

sense to the authors of this article is that we are in a better position at the end of a course to know what we did or did not know at the beginning of the course than we are when we are presented with an instrument that measures our self-efficacy for a course at the very beginning. We know more about what is included in the knowledge base of a course at the end than at the beginning. So, we are in a better place to appraise our ability at the end than at the beginning. A second finding of note is that students in the extension program class were less confident than the students in the main campus classes, even though their performance was not lower. This finding might shed some light on the influences upon self-efficacy that are outside the realm of performance.

A major limitation of this study was that it was focused upon only one category of the social work curriculum: research. Studies of other categories should be undertaken. Another limitation was that it was conducted in only one university.

The evaluation research self-efficacy tool employed in this study is rather general. Perhaps students are better able to appraise their ability when the instrument has very specific concepts. A specific example for research ability would be “the ability to find a statistical measure for testing the hypothesis when you are comparing the mean scores of two groups.” Another might be “to determine if a superior analysis of the target behavior in evaluative research is more useful for selecting a study sample, ~~or~~ selecting a means of measuring progress, or selecting an intervention.” The limitation of this method would be the enormous length of such an instrument if you-one wanted to really get take an in-depth look at the details of the course content. Sampling would be useful here. But if you-one wanted to find a way to measure performance, maybe perhaps it would be better more effective, and not more onerous, to just give

a sample of test questions rather than a sample of specific competencies in a specific self-efficacy instrument.

Commented [CM5]: This paragraph edited to match the tone and language of the rest of the article—OK?