DOCUMENT RESUME

ED 298 158 TM 012 263

AUTHOR Owen, Steven V.; Froman, Robin D.
TITLE Development of a College Academic Self-Efficacy Scale.
PUB DATE Apr 88
NOTE 8p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New Orleans, LA, April 6-8, 1988).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Aptitude; *College Students; Difficulty Level; Higher Education; *Rating Scales; Self Actualization; *Self Evaluation (Individuals); *Test Construction; Test Reliability; Test Validity
IDENTIFIERS *College Academic Self Efficacy Scale; *Likert Scales

ABSTRACT

This study concentrates on the development and estimation of measurement properties of the College Academic Self-Efficacy Scale (CASES). Three university faculties in education and psychology developed a pool of routine, frequent academic behaviors of college students. The pool was examined by seven graduate teaching assistants and trimmed and reworded based on their suggestions, before being rated by 93 undergraduate educational psychology students. A 33-item assortment without hierarchical arrangement resulted. A five-point Likert-type instruction was appended. For reliability estimation, the scale was administered twice over an eight-week period to another group of 88 educational psychology students. Concurrent validities were estimated with two different criteria, and factorial validity was estimated via exploratory principal factor analysis. Finally, a new sample was asked to estimate the difficulty of performing each of the behaviors listed in the scale. Results indicate that the preliminary measurement properties of the CASES are encouraging. Additional validity studies are recommended. (TJH)

a paper presented at NCME, New Orleans, April 1988



# Development of a College Academic Self–Efficacy Scale

Steven V. Owen, University of Connecticut
Robin D. Froman, University of Hartford





*Introduction.* Self–efficacy refers to a person's belief that he or she can carry out a particular behavior. Not long ago, Bandura (1977a) proposed self–efficacy as a good explanation for behavior change during psychotherapy. In the past decade, self–efficacy has traveled well: Not only is there increasing interest in the construct, but it is now attached to a sweep of behaviors far beyond its origins in anxiety and phobia. A quick check of *Psychological Abstracts* shows connections between self–efficacy and sports skills, academic performance, health practices, and socialization behaviors. Self–efficacy now occupies a central role in Bandura's (1986) social learning theory.

There is growing, sizable body of research suggesting that efficacy expectations influence motivation, persistence, and accomplishment. Nearly all studies, however, have focused on the content of self–efficacy rather than its measurement. Sometimes the measurements are so casual or confounded that one wonders what is actually being assessed. The present study concentrates on the development and estimation of measurement properties of a college–level academic self–efficacy scale.

*Instrument Background.* Bandura (1984) has stressed specificity in his self–efficacy scales, and items are usually arranged hierarchically (*e.g.*, lifting 10 pounds, lifting 20 pounds, etc.). The measurement often follows a two–step approach: First, a subject notes whether or not he can perform behavior A; then he indicates the "strength" of efficacy on a 90–point certainty scale. (Although Eastman and Marzillier [1984] have complained about using a 90––rather than 100––point scale, the practical difference is tiny.) The reason for the first step is to assess the "magnitude" of self–efficacy by summing the number of "can–do" assertions. An easier approach to measuring magnitude might be to indicate the level of task in the hierarchy where a person first claims "cannot–do." On the other hand, any measurement of magnitude may be flawed if "can–do" is modified by weak certainty in the second step.

Strength and magnitude *sound* like different aspects of self–efficacy, and theorists such as Cervone (1987) continue to promote their conceptual differences. A strictly empirical approach is to ask whether the two measures give unique and useful information about expectations or behavior. The small evidence in the literature shows that the two scales show substantial overlap. In their measure of academic self–efficacy, Wood and Locke (1987) showed correlations of .64 to .67 between strength and magnitude estimates. And these are spuriously low: the magnitude "scales suffered from a severe ceiling effect (and thus restriction of range)" (p. 1015). Wood and Locke's magnitude scales inconsistently explained unique variation in predicting academic goals and course performance, but the contribution was so small (averaging 2% of the criterion variance) that it has slim practical worth.

Interestingly, Wood and Locke concluded that the inconsistent results from magnitude measurement provide good reason to use "both [strength and magnitude] measures when assessing self–efficacy" (p. 1023).

Lee (1984) found a correlation of .83 between strength and magnitude scales of assertiveness. Fichten (1987) found a correlation of .95 between the measures and decided to drop the "strength" data from the rest of their research. In summary, strength and magnitude of efficacy beliefs are redundant, and if the analyst uses regression procedures, he invites garbled results because of multicollinear variables.

Further difficulties emerge because Bandura's Guttman–type hierarchical scales give spurious internal consistency estimates and factorial validities. Correlations approaching 1.00 between an item and preceding (easier) items seem guaranteed. Also, applying a hierarchical scale to cognitive behaviors may confuse the respondent. Wood and Locke (1987) developed a college academic self–efficacy measure in which they asked students about six behaviors, repeated across various percentages (termed "proportions" by Wood and Locke):

1. Memorize 60% of the facts & concepts [covered in this course]
2. Memorize 70% of the facts & concepts
3. Memorize 80% of the facts & concepts

etc. How are students able to distinguish among these levels? What frame of reference do students use for the 100% universe of course facts and concepts?

Other researchers have developed their own styles of measuring self–efficacy. Most of these measures are self–report, Likert–type scales. Reliability estimates are infrequently offered, and one wonders about the likely reliability of one–item (Newman & Goldfried, 1987) or 3–item scales (Maddux *et al.*, 1980). In a later paper, Maddux *et al.*, (1986) give a modest alpha estimate for their 3–item self–efficacy scale: .68. (In passing, we should acknowledge that some self–efficacy scales demonstrate good reliability [and validity] properties. Perhaps the most carefully developed and thoroughly studied is Wheeler and Ladd's [1982] Children's Self–Efficacy for Peer Interaction Scale [see also Perry *et al.*, 1986; Ladd & Price, 1986]).

Although Bandura (1986) has repeatedly stressed the difference between performance expectations (self–efficacy beliefs) and outcome expectations, some researchers disregard the distinction and build complex items that cover both beliefs. For example, "When a student gets a better grade than he usually gets, it is usually because I found better ways of teaching that student" (Gibson & Dembo, 1984). Others simply ignore the idea that self–efficacy beliefs refer to relatively circumscribed behaviors and develop massive questions: "How confident are you?", "How persistent are you?" (Englert & Tomlinson–Keasey, 1986); "If anyone can find the answer to what is troubling my baby, I am the one" (Cutrona & Troutman, 1986); "I feel that I am making a significant difference in the lives of my students" (Hoover–Dempsey *et al.*, 1987); "If parents would do more with their children, I could do more" (Gibson & Dembo, 1984). Still others use items that are obviously outcome beliefs and choose to call them self–efficacy questions (*e.g.*, Ashton & Webb, 1986).

Through the haze of available measurements, many of them offering weak or unknown psychometric properties, it seems unlikely that self–efficacy can penetrate clearly. For researchers who invest considerable energy in a project, it is tempting to

give dull data a little makeup. Ashton and Webb (1986, pp. 140–141), discussing the construct validity of their teacher self–efficacy measures, offer a correlation of .35 between alternate instruments as reasonable evidence. Or, they remark (p. 134) that "a relationship may be indicated" between a self–efficacy measure and teacher's emotional climate when the correlation approaches invisibility ($p$=.17). Later, they admit that the various scales "are inadequate from a psychometric perspective" and show "inadequate" internal consistency (p. 148). Thus, the self–efficacy literature is beginning to show the predictable blemishes of sloppy measurement: "Contrary to the theory...." and "In contrast to the findings of Smith...." The blemishes are more than cosmetic. Researchers, editors, reviewers, and readers untrained in measurement promote an ambiguous literature that can do great damage to a promising theory. They also can elevate worthless constructs to a high (if temporary) station.

These sorts of difficulties may lessen if a simpler and different measurement approach is taken. We tested this proposal in the development of the College Academic Self–Efficacy Scale (CASES).

*Methods and Results.* Three university faculty in education and psychology developed a pool of routine, frequent academic behaviors of college students. The pool was examined by 7 graduate teaching assistants, and was trimmed and reworded based on their suggestions. This revised pool was then given to 93 undergraduate educational psychology students. They were asked to rate each item on a 5–point, Likert–type scale, showing the "importance of the behavior in academic success." Items with mean importance ratings of less than 3.0 (where 5 = extremely important) were dropped, leaving a 33–item assortment without hierarchical arrangement. The revised items range from very specific (attending class consistently in a dull course) to fairly general (understanding difficult passages in textbooks). A 5–point Likert–type instruction was now appended: How much confidence do you have about performing each behavior listed below? Only the poles (Very Little...Quite A Lot) were labeled. Figure 1 illustrates the layout of the questionnaire.

---

Insert Figure 1 about here

---

For reliability estimation, the scale was administered twice over an 8–week interval to another group of 88 educational psychology students. Alpha internal consistency estimates for the two occasions were .90 and .92. The 8–week stability estimate was .85

Concurrent validities were estimated with two different criteria, each suggested by self–efficacy theory: *frequency* of performing each task, and *enjoyment* of each task. In separate studies, students were asked for 5–point self–ratings on frequency and enjoyment for each of the 33 academic behaviors on CASES. These studies were arranged as incremental validity research. In predicting mean item *frequency*, grade–point average (GPA) was forced into the regression equation, followed by the CASES score. An identical analysis was done for the prediction of mean item *enjoyment*. The analyses, each carried out on different samples of educational psychology students from those described above, gave very similar results:

| *Criterion* | *Predictor* | *R* | *N* |
|---|---|---|---|
| frequency | 1. GPA | .40 | 127 |
| | 2. CASES | .78 | |
| enjoyment | 1. GPA | .38 | 79 |
| | 2. CASES | .72 | |

In short, academic self–efficacy showed very strong incremental validity beyond that explained by GPA alone.

In a variation of these concurrent validity studies, the two samples were combined, and course grade was regressed hierarchically on GPA, then CASES score. The addition of CASES increased *R* from .62 to .81.

To study factorial validity, responses from three samples were combined, and an exploratory principal factor analysis was performed. Three factors with eigenvalues above 1.0, explaining 78% of the systematic item variance, were sent to an orthogonal rotation. A clear and easily interpretable structure emerged. The three factors were called (1) Overt, Social Situations (*e.g.*, participating in a class discussion), (2) Cognitive Operations (*e.g.*, listening carefully during a lecture on a difficult topic), and (3) Technical Skills (*e.g.*, using a computer).

Finally, a new sample ($n = 122$) was asked to estimate the difficulty of performing each of the behaviors listed in CASES for most students. Mean item difficulties were calculated and inspected for general confirmation of self–efficacy theory. We found that the behaviors rated least difficult were ones that students probably had had the most experience––and success––with (examples: attending classes regularly, performing simple math computations). And the behaviors rated most difficult were those students rarely try, rarely succeed at, and rarely observe models succeeding at (examples: challenging a professor's opinion in class, answering a question in a large class). This intuitive analysis seems to confirm predictions made by self–efficacy theory (Bandura, 1986).

*Educational Importance.* The preliminary measurement properties of CASES are encouraging. In addition, CASES is a rapid (5 minutes) and simple tool for assessing academic self–efficacy. We are now testing a very practical application of the instrument. Because those with low academic self–efficacy may be at risk for probation or dropping out, selected students are targeted for academic counseling to prevent failure. In fact, certain CASES items are specific enough to give explicit diagnostic direction (for example, Spreading studying out instead of cramming).

Additional validity studies are being prepared. Objective criterion measures––as opposed to self–report––are being collected. We are studying how to modify the scale to accommodate students majoring in unusual areas, such as art or music. We are also planning a study of how social desirabilty may influence self–report efficacy scores. Finally, we are investigating whether item weighting may improve CASES' measurement properties.

*References*


Ashton, P.T. & Webb, R.B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement.* NY: Longman.

Bandura, A. (1977). Self–efficacy: Toward a unifying theory of behavioral change. *Psychol. Rev., 84,,* 191–215.

Bandura, A. (1984). Recycling misconceptions of perceived self–efficacy. *Cogn. Ther. & Res., 8,,* 231–255.

Bandura, A. (1986). *Social foundations of thought and action.* Englewood Cliffs, NJ: Prentice–Hall.

Cervone, D. (1987). Chi–square analyses of self–efficacy data: A cautionary note. *Cogn. Ther. & Res., 11*(6), 709–714.

Cutrona, C.E. & Troutman, B.R. (1986). Social support, infant temperament, and parenting self–efficacy: A mediational model of postpartum depression. *Child Devel., 57,* 1507–1518.

Eastman, C. & Marzillier, J.S. (1984). Theoretical and methodological difficulties in Bandura's self–efficacy theory. *Cogn. Ther. & Res., 8,,* 23–230.

Englert, A.M. & Tomlinson–Keasey, C. (1986). Competence and self–efficacy among the gifted. A paper presented at the annual meeting of the Americal Psychological Assn., Washington, D.C.

Fichten, C.S., Bourdon, C.V., Amsel, R., & Fox, L. (1987). Validation of the college interaction self–efficacy questionnaire: Students with and without disabilities. *J. of Coll. Student Personnel., 29,* 449–458.

Gibson, S. & Dembo, M.H. (1984). Teacher efficacy: A construct validation. *J. Educ. Psychol., 76*(4), 569–582.

Hoover–Dempsey, K.V., Bassler, O.C., & Brissie, J.S. (1987). Parent involvement: Contributions of teacher efficacy, school socioeconomic status, and other school characteristics. *Amer. Educ. Res. J., 24* 417-435.

Ladd, G. & Price, J.M. (1986). Promoting children's cognitive and social competence: The relation between parents' perceptions of task difficulty and children's perceived and actual competence. *Child Devel., 57,* 446–460.

Lee, C. (1984). Accuracy of efficacy and outcome expectations in predicting performance in a simulated assertiveness task. *Cogn. Ther. & Res., 8*(1), 37–48.

Maddux, J.E., Norton, L.W., & Stoltenberg, C.D. (1986). Self–efficacy expectancy, outcome expectancy, and outcome value: Relative effects on behavioral intentions. *J. of Person. & Soc. Psychol., 51*(4), 783–789.

Maddux, J.E., Sherer, M., & Rogers, R.W. (1980). Self–efficacy expectancy and outcome expectancy: Their relationship and their effects on behavioral intentions. A paper presented at the annual meeting of the Southeastern Psychological Assn., Washington, D.C.

Newman, C. & Goldfried, M.R. (1987). Disabusing negative self–efficacy expectations via experience, feedback, and discrediting. *Cogn. Ther. & Res., 11*(4), 401–417.

Perry, D.G., Perry, L.C., & Rasmussen, P. (1986). Cognitive social learning mediators of aggression. *Child Devel., 57,* 700–711.

Wheeler, V.A. & Ladd, G.W. (1982). Assessment of children's self–efficacy for social interactions with peers. *Devel. Psychol., 18*(6), 795–805.

Wood, R.E. & Locke, E.A. (1987). The relation of self–efficacy and grade

goals to academic performance. *Educ. & Psychol. Meas.*, *47*, 1013–1024.

How much confidence do you have about doing each of the behaviors listed below? Circle the letters that best represent your beliefs.

| A | B | C | D | E |
|---|---|---|---|---|
| very little | <---------- | *CONFIDENCE* | ----------> | quite a lot |

A B C D E 1. Taking well organized notes during a lecture.
A B C D E 2. Participating in a class discussion.
A B C D E 3. Answering a question in a large class.

Figure 1. Example items from CASES.