Reliability and Validity of the Physical Self-Efficacy Scale in a Competitive Sport Setting

By: Edward McAuley and Diane Gill

McAuley, E. & Gill, D.L. (1983). Reliability and validity of the Physical Self-Efficacy Scale in a competitive sport setting. Journal of Sport & Exercise Psychology, 5, 410-418.

*****Note: Figures may be missing from this format of the document**

Abstract:
Interest in the role of self-confidence in sport performance has been high in sport psychology research. A measure to assess general physical self-efficacy has recently been developed, but without application to competitive sport performance. The present study examined the role of general and task-specific self-efficacy in women's intercollegiate gymnastics. It also assessed the reliability and validity of the Physical Self-Efficacy Scale in a competitive sport setting. The Physical Self-Efficacy Scale was found to be a reliable and valid instrument for measuring an individual's general physical self-efficacy in sport. However, the task-specific measures of self-efficacy and the gymnast's prediction of how they would perform proved to be much more powerful variables for predicting actual gymnastic performance. The results are discussed in terms of the relationships between different types of self-efficacy and sport performance and the problems associated with self-efficacy measurement.

Article:
Athletes, coaches, and spectators regard self-confidence as a necessary quality for successful sport performance. However, a highly confident basketball player may feel comfortable on a basketball court, but exhibit very low self-confidence if required to perform a routine on a trampoline. Similarly, that same athlete may be very confident dunking the basketball but not at all confident dribbling the ball up the court during a game. Self-confidence, therefore, appears to be a very situational- specific quality. Bandura (1977) has labeled this type of specific self-confidence as self-efficacy and purports it to be the common cognitive mechanism mediating behavioral change. Self-efficacy expectations influence persistence, thought patterns, arousal, and ultimately behavior (Bandura, 1977; Bandura, Adams, & Beyer, 1977; Bandura & Schunk, 1981). Perceived self-efficacy is thought to determine behavioral outcomes when sufficient incentives and the required skills are present. Support for self-efficacy as a theory of behavioral change has been presented by Bandura and his colleagues as well as a number of researchers in the area of sport psychology (e.g., Feltz, 1982; Feltz & Mugno, 1983; Feltz, Landers, & Raeder, 1979; Weinberg, Gould, Yukelson, & Jackson, 1981).

As self-efficacy is a situation-specific construct, measures which adequately assess self-efficacy for different behavioral domains are needed. Given the myriad situations in which self-efficacy is of concern, attempts to develop countless scales for every conceivable situation are unrealistic. Construction of scales for general areas of behavior is a more viable proposition.

To date, self-efficacy has been assessed by numerous researchers following the recommendations of Bandura and his colleagues (1977). Specifically, respondents are asked to indicate how many items on a behavioral checklist they can successfully complete. The number of behaviors indicated is considered to be the level of efficacy. The respondent also indicates how confident he/she is that he/she can complete each behavior on a 100-point percentage scale. The confidence ratings are summed and divided by the total number of items to create a strength of self-efficacy score (e.g., a total confidence score of $300^{0}10$ for six items on a 10-item scale would yield an efficacy level of six and a strength of efficacy of 300/10 = 30). The development of a more general efficacy measure for sport or physical activity would allow researchers to explore the importance of general self-efficacy in successful sport or athletic performance.

Ryckman, Robbins, Thornton, and Cantrell (1982) have recently developed the Physical Self-Efficacy Scale (PSE), a measure of an individual's perceived physical self-confidence. The scale is constructed with two subscales, the Physical Self-Presentation Confidence scale (PSPC) and the Perceived Physical Ability scale (PPA). Ryckman et al. reported adequate reliability and validity for the PSE and its subscales. Of particular interest to this investigation is the reported predictive validity of the PPA. Individuals with a high score on the PPA were reported to perform better and more accurately predict performance on a motor skills task than individuals scoring low on the PPA. Ryckman and his colleagues suggest that general physical self-efficacy is related to performance of simple motor tasks. However, it is unclear as to the role physical self-efficacy plays in the performance of more complex physical activities such as sport and athletics. It would be logical to propose that physical self-efficacy influences a more task-specific self-efficacy which, in turn, affects how well one expects to perform (i.e., prediction of performance), which ultimately influences performance. This causal model is slightly more complex than the relationships between physical self-efficacy, prediction, and performance that are proposed by Ryckman et al. and, therefore, needs to be tested.

Ryckman and his colleagues have stated that the PSE scale, and in particular the PPA, has relevance for some instances in physical education and athletic programs. The present study focuses upon intercollegiate athletics, specifically women's gymnastics. The purpose was to compare the PSE to more task-specific measures of self-efficacy and to explore the relationships between these measures and performance. The reliability and validity of the PSE in an athletic situation was also assessed.

Method

Subjects

Female collegiate gymnasts *(N* = 52) representing seven universities volunteered for the present study. The head coaches at each of the universities granted permission, by letter or telephone, for the investigators to approach the gymnasts immediately prior to official timed warm-ups. The gymnasts completed informed consent forms and then completed the inventory assessing their confidence going into the meet.

Table 1

Items Used in Situation-Specific Self-Efficacy Measures

Vault	Beam
Handspring	Back handspring
Half on half off	Aerial cartwheel
Handspring full twist	Aerial walkover
Tsukahara (tuck or pike)	Double turn on one foot
Half on full off	Back tuck salto
Layout Tsukahara	Punch front salto
Handspring front	Back layout stepout
Floor	**Bars**
Back salto	Cut catch
Front salto	Peach
Arabian	Front salto
Layout full twist	Cast wrap low bar full
Triple turn on one foot	Cast handstand straight
Double full twist	Free hip to handstand
Double back salto	Giant swing to handstand

Procedure and Dependent Measures

Each gymnast was asked to complete a self-confidence inventory prior to beginning official warm-ups. The inventory consisted of the Physical Self-Efficacy Scale (PSE) and four task-specific efficacy inventories. These task-specific measures were constructed by two Class I U.S.G.F. judges and comprised seven gymnastic elements listed in order of ascending difficulty. One measure was constructed for each of the four Olympic events: vault, bars, balance beam, and floor exercise (see Table 1). The gymnasts were asked to indicate how many of the items on each scale that they thought they could successfully complete at that point in time and also how confident they were that they could complete each item. The gymnasts also predicted their actual score on each event that they were competing in that day. Administration of the PSE and the task-specific measures of self-efficacy was counterbalanced to control for possible order effects. At the conclusion of the gymnastics meet and following the awards ceremony, the meet director provided the investigators with a copy of the official scoresheet which detailed individual scores on each event.

Results

The data were analyzed in a number of stages. Analyses were conducted to determine the factor structure, the reliability, and the validity of the Physical Self- Efficacy (PSE) Scale in an applied sport setting.

Factor Analysis

Ryckman et al. (1982) present evidence to support a two-factor structure for the PSE. Principal components analyses revealed two principal loadings representing a Perceived Physical Ability (PPA) dimension and a Physical Self-Presentation Confidence (PSPC) dimension. The two factors comprise a total of 22 items from an original pool of 90 items administered to a sample of 363 subjects. To examine the accuracy of this factor structure a confirmatory factor analysis was implemented using the Lisrel V program (Joreskog & Sorbom, 1981). This program employs a measurement model to test the hypothetical factor structure. Using least squares estimates and uncorrelated error specifications the analysis revealed an overall goodness of fit index of .65 for the hypothetical two-factor model, $X^2(208) = 319.4$, $p < .0001$. The magnitude of this statistic, however, is misleading and Joreskog (1969) suggests that the ratio between the chi-square and its degrees of freedom is a better means of evaluating how well a model fits the data. The ratio of the chi-square to the degrees of freedom suggested is 2:1. Following such guidelines, the two-factor structure does appear to be adequate. The relatively low goodness of fit index can be attributed to the small number of subjects in the present study *(N* = 52).

Reliability

The internal consistency of the Physical Self-Efficacy scale and its two underlying dimensions were assessed via coefficient alpha (Cronbach, 1951). Coefficient alpha for the 10 items of the PPA had a value of .76, suggesting that this factor was being reliably assessed. The other factor, PSPC, had an internal consistency value of .42 over 12 items and appears less reliable in the present study. Coefficient alpha for the overall PSE scale was adequate with a value of .72. The PSPC does appear to have some reliability problems in this particular sample and researchers should be cognizant of this weakness if they intend to use this particular subscale. However, for the purpose of the present study, in which the interest is primarily in the predictive powers of the overall scale and the PPA, the reliability coefficients are acceptable.

Validity

If the PSE scale does accurately measure an individual's perceived physical self-efficacy, criterion validity can be demonstrated by examining the relationship between the PSE and other measures of self-efficacy. Correlation analyses were conducted to examine these relationships. On vault, the PSE and vaulting self-efficacy were positively related, $r = .26$, $p < .05$. On uneven bars, the PSPC and the PPA were positively related to the bars self-efficacy, as was the overall PSE, $r = .40$, $p < .005$. On balance beam, the PSE and the PSPC correlated positively with beam

Table 2

Correlations Among the PSE Scales and Self-Efficacy by Event

	Vault	Beam	Bars	Floor exercise
PSE	.258*	.297*	.401***	.359**
PPA	.191	.119	.331**	.359**
PSPC	.248	.392**	.328**	.311*

***$p < .001$
**$p < .01$
*$p < .05$

Table 3

Correlations Between Performance Scores and Efficacy Measures

	Vault score	Beam score	Bars score	Floor score
PPA	−.033	.181	.264*	.199
PSPC	.067	.197	.166	.252
PSE	.016	.223	.257*	.286*
Vault efficacy	.284*			
Beam efficacy		.582***		
Bars efficacy			.717***	
Floor efficacy				.426**

***$p < .001$
**$p < .01$
*$p < .05$

self-efficacy. On floor exercise, both subscales and the total PSE correlated positively with the specific self-efficacy measure. All correlations and significance levels are shown in Table 2. It appears that the PSE and its subscales do exhibit some criterion validity in measuring perceptions of general physical self-efficacy. However, one must consider this relationship between general self-efficacy and the task-specific self-efficacy measures with caution as the two measures share only 10°7o of the variance in common.

The predictive validity of the PSE was examined next. According to the developers (Ryckman et al., 1982), the PSE should be related to the criterion variable (performance score on each event). It is logical to assume that this relationship will be weaker than the relationship between performance and task-specific measures of self-efficacy. These relationships are shown in Table 3. The PSE correlated significantly with scores on bars and floor exercise, but not with scores on vault and beam. On the other hand, all task-specific measures of self-efficacy correlated significantly with performance scores in their respective events (see Table 3) accounting for up to 52% of the variance in performance score. Furthermore, the gymnasts' own predictions of performance scores correlated more strongly with actual scores than the task-specific measures.

Thus far, the analyses suggest that although a weak relationship exists between the PSE and more task-specific measures of efficacy expectations, the general measure is an inadequate predictor of the performance criterion. Specific measures of self-efficacy, and in particular the actual predictions of performance made by the gymnasts, were much more accurate predictors of performance.

The role played by physical self-efficacy in gymnastic performance was established by examining a causal model that views performance as a function of physical self-efficacy, plus specific self-efficacy, plus the gymnast's prediction of performance. In path analytical terms given the results so far, direct paths should be present between prediction and performance, task-specific self-efficacy and performance, and physical self-

efficacy and specific self-efficacy. Path analysis allows assessment of the importance of the performance predictions after controlling for other variables in the model (PSE and specific efficacy expectations). To examine the suggested causal model hierarchical multiple regression analyses were conducted. One set of analyses used the performance score on each event as the criterion (Y) variable and the PSE, task-specific efficacy, and the performance predictions as the predictor (X) variables. As Ryckman et al. (1982) postulated that the PPA scale is a significant predictor of performance of motor skill tasks, a second set of multiple regression analyses was conducted using the same procedures, but replacing the PSE with the PPA scale in the regression equations.

No differences in results were observed between the two sets of analyses using the PPA versus the PSE as predictors. Therefore, only the results in which the PSE was used are presented. It should be noted that when the predictor variables in a causal model are intercorrelated, the beta weights in the final regression equation are reduced and path diagrams including these beta weights render little information of value. More appropriate and of more value is the reporting of the R-squared change with the addition of each predictor variable and an F-test to assess the statistical significance of the change in percent variance. Table 4 presents all of this relevant information.

As Table 4 indicates, the PSE was not a significant predictor of performance score in any of the four events, accounting for between .0001% and .07% of the variance in performance. Task-specific self-efficacy, on the other hand, was a significant predictor of performance in three of the events and approached significance in the vault. Of added interest is the finding that the gymnasts' predictions of their own performance scores in each event accounted for a significant proportion of variance in the performance score over and above that accounted for by task-specific self-efficacy. The individual's global prediction of performance may well be a summary of that person's efficacy expectations as well as physical and mental preparation for competition. If the predicted scores indeed include the gymnast's perceptions of her self-efficacy on each event, then task-specific efficacy should not be a significant predictor of performance when the gymnasts' predictions of performance are statistically controlled. A further set of multiple regression analyses were computed to examine the accuracy of this assumption. As expected, the self-efficacy variable was indeed a nonsignificant predictor when entered last into the regression equation. In summary, it appears that the PSE, while being weakly related to specific

Table 4

Multiple Regression Analyses by Event for Performance Score

Predictor	Beta	R^2	R^2 change	F for R^2 change
PSE	−.09	.066	.066	2.62
Efficacy beam	.15	.349	.283	15.70***
Predicted score	.71	.608	.259	23.18***
PSE	−.07	.063	.063	2.38
Efficacy floor	.08	.187	.123	5.19***
Predicted score	.76	.605	.417	34.93***
PSE	.07	.066	.066	2.69
Efficacy bars	.24	.514	.448	34.18***
Predicted score	.62	.656	.141	14.80***
PSE	−.14	.000	.000	.00
Efficacy vault	.05	.087	.087	3.61*
Predicted score	.68	.411	.324	20.43***

***$p < .001$
**$p < .01$
*$p < .1$

self-efficacy, is not a significant predictor of motor performance as stated by Ryckman and his associates (1982).

Discussion

The purpose of the present study was to examine the reliability and validity of the Physical Self-Efficacy Scale (Ryckman et al., 1982) in a competitive sport setting and examine the role of physical and task-specific self-efficacy in gymnastic performance. Data were collected from 52 female collegiate gymnasts prior to a seven school invitational meet. The PSE, according to the developers of the scale, is a two factor-scale which assesses an individual's physical self-confidence across situations. Initial analyses examined the factor structure proposed by Ryckman et al. (1982). Considering the small sample population used in this study *(N* = 52), the hypothetical factor structure appears quite tenable. The two factors examined in the confirmatory factor analysis were the Physical Self-Presentation Confidence (PSPC) and the Perceived Physical Ability (PPA). The assessment of the internal consistencies of the two subscales and the overall scale, using coefficient alpha, demonstrated adequate reliability for the overall scale and the PPA. The PSPC, however, had a rather low alpha coefficient (.42) and a note of caution should be administered to those researchers who intend to use the PSPC. From the results of the present study one can conclude that the factor structure and the reliability of the Physical Self- Efficacy Scale are adequate.

Criterion validity for the PSE and its subscales was demonstrated by examining the relationships between the PSE and task-specific measures of self-efficacy. The PSE and its subscales correlated significantly with these measures demonstrating criterion validity. Of course, one would not want to assume that the task-specific measures are completely accurate and valid criterion measures. For example, a gymnastic routine is composed of a large number of elements, which differ from individual to individual because the composition of routines at the collegiate level is optional. Thus, the measure may be accurate for one individual but not for another. In constructing the measures an attempt was made to have an adequate cross section of skills represented.

Predictive validity was assessed by examining the relationship between performance scores on each of the events and the PSE. These relationships were nonsignificant. Given that the PSE is a measure of general physical self-efficacy over situations, one would not expect a strong relationship between the PSE and actual performance on the criterion variable (performance score). However, task-specific measures of self-efficacy and the gymnasts' predictions of performance scores correlated moderately with performance scores. These measures, predicted score and self-efficacy, were highly intercorrelated and obviously are not orthogonal constructs. The individual's knowledge, experience, and past accomplishments (all sources of efficacy information) apparently combine to form a more accurate representation of event-specific efficacy expectations than do measures constructed by researchers, judges, and coaches. However, as previous research suggests, task-specific self-efficacy appears to be an important determinant of performance.

Regression analyses were used to assess differential effects of the PSE, the task-specific measures of self-efficacy, and the predicted score for each event on actual perfomance score. The PSE accounted for very little of the variance in performance score, but the task-specific efficacy measures and predicted scores accounted for substantial amounts of the variance in performance. Of further interest was the fact that predicted scores were significant predictors of performance even when the PSE and self-efficacy predictor variables were statistically controlled. These results suggest that the individual may use other information sources in combination with self-efficacy expectations to assess how well he/she is going to perform in competition. These data suggest that future research in the area of perceived self-efficacy and its mediating effects upon athletic performance should employ task-specific measures of self-efficacy rather than more general measures such as the Physical Self-Efficacy Scale. These findings need to be replicated in more diverse athletic settings and with larger sample sizes. A further note of caution is warranted. The present study examined the role of self-efficacy in women's gymnastics and generalization of the results to men's gymnastics cannot be recommended without further research employing male subjects.

In summary, the PSE appears to be a reliable and, to a certain extent, valid measure of general physical self-efficacy. Contrary to the suggestions of the developers of the scale, the PSE was not found to be a significant predictor of performance in sport skills. A weak relationship was found between the PSE and task- specific self-efficacy and predictions of performance, which were in turn significant predictors of actual performance.

References

Bandura, A. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 1977, 84, 191-215.

Bandura, A., Adams, N.E., & Beyer, J. Cognitive processes mediating behavioral change. *Journal of Personality and Social Psychology*, 1977, 35, 125-139.

Bandura, A., & Schunk, D.H. Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology*, 1981, 586-598.

Cronbach, L.J. Coefficient alpha and the internal structure of tests. *Psychometrika*, 1951, 16, 296-334.

Feltz, D.L. Path analysis of the causal elements in Bandura's theory of self-efficacy and an anxiety-based model of avoidance behavior. *Journal of Personality and Social Psychology*, 1982, 42, 764-781.

Feltz, D.L., & Mugno, D.A. A replication of the path analysis of the causal elements in Bandura's theory of self-efficacy and the influence of autonomic perception. *Journal of Sport Psychology*, 1983, 5, 263-277.

Feltz, D.L., Landers, D.M., & Raeder, U. Enhancing self-efficacy in high avoidance motor- tasks: A comparison of modeling techniques. *Journal of Sport Psychology*, 1979, 1, 112-122.

Joreskog, K.G. A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 1969, 34, 183-202.

Joreskog, K.G., & Sorbom, D. *Lisrel V: Analysis of linear structural relationships by the method of maximum likelihood.* Chicago: National Education Resources, Inc., 1981.

Ryckman, R.M., Robbins, M.A., Thornton, B., & Cantrell, P. Development and validation of a Physical Self-Efficacy Scale. *Journal of Personality and Social Psychology*, 1982, 891-900.

Weinberg, R.S., Gould, D., Yukelson, D., & Jackson, A. The effect of preexisting and manipulated self-efficacy on a competitive muscular endurance task. *Journal of Sport Psychology*, 1981, 4, 345-354.