European Journal of Psychological Assessment, Vol. 18, Issue 3, pp. 242-251

Is General Self-Efficacy a Universal Construct?*

Psychometric Findings from 25 Countries

Urte Scholz[1], Benicio Gutiérrez Doña[2], Shonali Sud[3], and Ralf Schwarzer[1]

[1]Freie Universität Berlin, Germany, [2]Universidad Estatal a Distancia, Costa Rica,
[3]St. Bede's College, Shimla, India

Keywords: Cross-cultural research, self-efficacy, optimism, psychometric scale adaptation

Summary: Perceived self-efficacy represents an optimistic sense of personal competence that seems to be a pervasive phenomenon accounting for motivation and accomplishments in human beings. The General Self-Efficacy scale, developed to measure this construct at the broadest level, has been adapted to many languages. The psychometric properties of this instrument is examined among 19,120 participants from 25 countries. The main research question is whether the measure is configurally equivalent across cultures, that is, whether it corresponds to only one dimension. The findings confirm this assumption and suggest the globality of the underlying construct. They also point to a number of cross-cultural differences that merit further investigation.

Introduction

The Construct of Perceived Self-Efficacy

Self-referent thought has become an issue that pervades psychological research in many domains. It has been found that a strong sense of personal efficacy is related to better health, higher achievement, and better social integration (Schwarzer, 1992; Bandura, 1997). The construct of self-efficacy represents one core aspect of Bandura's social-cognitive theory (Bandura, 1977, 1997, 2000, 2001). Bandura, in a unifying theory of behavior change, hypothesized that expectations of self-efficacy determine whether coping behavior will be initiated, how much effort will be expended, and how long it will be sustained in the face of obstacles and aversive experiences. While outcome expectancies pertain to the perception of possible consequences of one's action, perceived self-efficacy refers to personal action control or agency. A person who believes in being able to produce a desired effect can lead a more active and self-determined life. This "can do"-cognition mirrors a sense of control over one's environment. It reflects the belief of being able to control challenging environmental demands by taking adaptive action. It can be regarded as an optimistic and self-confident view of one's capability to deal with certain life stressors.

According to theory and research, self-efficacy makes a difference in how people feel, think and act (Bandura, 1997). In terms of feeling, a low sense of self-efficacy is associated with depression, anxiety, and helplessness. Persons with low self-efficacy also have low self-esteem, and they harbor pessimistic thoughts about their accomplishments and personal development. In terms of thinking, a strong sense of competence facilitates cognitive processes and performance in a variety of settings, including quality of decision-making and academic achievement. Self-efficacy has an influence on preparing action because self-related cognitions are a major ingredient in the motivation process. Self-efficacy levels can enhance or impede motivation. People with high self-efficacy choose to perform more challenging tasks (Bandura, 1997). They set themselves higher goals and stick to them. Actions are preshaped in thought, and people anticipate either optimistic or pessimistic scenarios in line with their level of self-efficacy. Once an action has

* The original data upon which this paper is based are available at http://www.hhpub.com/journals/ejpa

EJPA 18 (3), © 2002 Hogrefe & Huber Publishers
DOI: 10.1027//1015-5759.18.3.242

been taken, highly self-efficacious people invest more effort and persist longer than those low in self-efficacy. When setbacks occur, they recover more quickly and maintain commitment to their goals. High self-efficacy also allows people to select challenging settings, explore their environment, or create new ones.

According to Bandura, there are four major sources for influencing personal competence. First, self-efficacy beliefs can be enhanced through *personal accomplishment* or *mastery*, as far as success is attributed internally and can be repeated. A second source is *vicarious experience*. When a "model person" who is similar to the individual successfully masters a difficult situation, social comparison processes can enhance self-efficacy beliefs. Third, there is *symbolic experience* through verbal persuasion by others (e. g., a teacher reassures a student that she will certainly pass the exam due to her academic competence). The last source of influence is *emotional arousal*, that is, the person experiences anxiety in a threatening situation and thus feels incapable of mastering the situation. These four informational sources vary in strength and importance in the order presented here.

Self-efficacy is commonly understood as being domain-specific. That is, one can have more or less firm self-beliefs in different domains or particular situations of functioning. But some researchers have also conceptualized a generalized sense of self-efficacy that refers to global confidence in one's coping ability across a wide range of demanding or novel situations (Sherer & Maddux, 1982; Skinner et al., 1988; Schwarzer & Jerusalem, 1999). General self-efficacy (GSE) aims at a broad and stable sense of personal competence to deal effectively with a variety of stressful situations (Schwarzer, 1992; Schwarzer et al., 1999). The present authors agree with Bandura (1997) that perceived self-efficacy should be conceptualized in a situation-specific manner. However, the degree of specificity of generality varies with the context. If the research question deals with solving an algebra problem or running a marathon, the wording of self-efficacy items will be more narrow than when the professional self-efficacy of teachers or nurses is at stake (Schwarzer & Jerusalem, 1999; Schwarzer et al., 2000).

The highest level of generality is given when broad optimistic self-beliefs are examined, for example, when individuals under stress have to readapt to novel life circumstances over an extended period of time. In a study with cardiac surgery patients, Schröder et al. (1998) found that patients with high GSE scores had recovered better one week after surgery and experienced better quality of life half a year later than their low-GSE counterparts. In a study among East German refugees, people with high GSE were healthier, socially better integrated, and more frequently employed two years after the stressful transition than their low-GSE counterparts (Schwarzer et al., 1993).

Research Question

The purpose of the present study is to examine the psychometric properties of the General Self-Efficacy Scale in 25 samples. Item-level statistics and analyses of the internal structure of the instrument, including reliability, are conducted. Principal component analyses (PCA) and confirmatory factor analyses (CFA) test a one-factor against a multifactor solution in order to corroborate the unidimensionality of the construct across nations. Furthermore, mean differences between nation and gender are analyzed.

Method

The General Self-Efficacy Scale

The German version of the General Self-Efficacy (GSE) scale was originally developed by Matthias Jerusalem and Ralf Schwarzer in 1979. This instrument contained 20 items. In 1981 it was reduced to 10 items and subsequently adapted to 28 languages (see below; Schwarzer & Jerusalem, 1995). A typical item is, "Thanks to my resourcefulness, I can handle unforeseen situations" (see Appendix).

The GSE scale has been used in numerous research projects, where it typically yielded internal consistencies between alpha = .75 and .91. Its stability has been examined in several longitudinal studies. For example, 246 cardiac surgery patients in Germany who filled out the questionnaire before surgery and again half a year later, had a retest-reliability of $r = .67$ (Schröder et al., 1998). Among 140 teachers in Germany, a stability coefficient of $r = .75$ was found after one year. Over the same time period, 2846 students, also in Germany, filled out the scale twice, whereby a retest reliability of $r = .55$ was found (Schwarzer & Jerusalem, 1999). In addition, for a two-year period, there were coefficients of $r = .47$ for men and $r = .63$ for women who had left East Germany as refugees (Schwarzer et al., 1993).

A large-scale German field research project with 3514 high-school students and 302 teachers has provided evidence for validity of the GSE scale (Schwarzer & Jerusalem, 1999). For the group of students, general self-efficacy correlated .49 with optimism and .45 with the perception of challenge in stressful situations. For the teachers, high correlations were obtained with proactive coping (.55), self-regulation (.58), and procrastination

Table 1. Number of participants and mean age broken down by country and gender.

	Men	Women	Missing	Total	Mean Age Men	Mean Age Women	Mean Age Total
Belgium	–	–	204	204	–	–	–
Canada	180	187	–	367	15.2	15.1	15.1
Costa Rica	356	607	7	970	21.0	21.3	21.2
Denmark	–	–	163	163	–	–	–
Finland	–	–	163	163	–	–	–
France	–	–	144	144	–	–	–
Germany	3309	3313	801	7423	22.6	25.8	24.3
Great Britain	26	193	242	461	57.5	60.8	60.4
Greece	50	50	–	100	42.6	38.2	40.4
Hong Kong	342	724	1	1067	16.4	16.4	16.4
Hungary	25	39	95	159	20.8	22.3	23.9
India	181	217	–	398	19.9	19.5	19.6
Indonesia	276	260	–	536	21.1	20.4	20.8
Iran	415	383	4	802	22.8	21.7	22.3
Italy	–	–	148	148	–	–	–
Japan	194	236	–	430	19.1	18.5	18.8
Korea	35	111	1	147	67.1	62.3	63.2
Netherlands	185	537	226	948	59.0	61.4	60.8
Peru	371	658	–	1029	19.0	18.8	18.9
Poland	155	415	127	697	26.0	25.9	25.9
Portugal	211	238	119	568	15.4	15.2	15.3
Russia	205	290	–	495	25.9	28.4	27.4
Spain	–	–	429	429	–	–	–
Syria	115	149	–	264	28.3	24.6	26.2
USA	823	793	17	1633	–	–	–

(–.56). Moreover, there was a substantial relationship to all three dimensions of teacher burnout (emotional exhaustion –.47, depersonalization –.44, and lack of accomplishment –.75). Similar evidence for validity was found for teachers in Hong Kong (Schwarzer et al., 2000).

Bilingual native speakers adapted the ten self-efficacy items to 28 languages, based on the German and English versions of the GSE scale. The adaptations followed the "group consensus model" with several bilingual translators. The procedure included back translations and group discussions. Since the goal was to achieve cultural-sensitive adaptations of the construct rather than mere literal translations, the translators acquired a thorough understanding of the general self-efficacy construct. To date, several studies have been published that compare the psychometric properties of the German, English, Dutch, Spanish, Russian, Greek, Arabic, Hungarian, Polish,

Table 2. Item means and corrected item-total correlations for each nation (*n* in parentheses).

	Belgium (175)		Canada (367)		Costa Rica (943)		Denmark (153)		Finland (159)		France (103)		Germany (7100)		Great Britain (447)	
Item	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)
1	3.16	.49	3.19	.62	3.09	.25	3.42	.62	2.94	.56	3.23	.58	3.12	.42	3.20	.54
2	2.39	.34	3.01	.36	3.75	.36	2.72	.31	1.89	.32	2.80	.43	3.08	.43	2.44	.35
3	2.93	.24	2.97	.56	3.17	.35	3.17	.64	2.74	.33	3.30	.39	2.88	.46	2.87	.57
4	3.16	.67	3.10	.65	3.32	.56	3.33	.71	3.04	.66	3.40	.65	2.78	.50	3.11	.72
5	3.26	.62	3.16	.68	3.21	.61	3.44	.63	3.11	.65	3.34	.57	2.95	.49	3.06	.73
6	3.27	.61	3.37	.62	3.34	.63	3.56	.63	3.31	.58	2.94	.30	2.75	.54	3.23	.69
7	3.18	.66	3.04	.60	3.22	.64	3.36	.66	3.06	.64	3.44	.60	3.02	.54	3.14	.67
8	2.99	.55	3.00	.70	3.65	.51	3.14	.53	2.74	.52	3.16	.59	2.98	.51	2.90	.58
9	3.12	.64	3.07	.56	3.09	.54	3.49	.53	2.92	.68	3.22	.47	3.02	.54	3.10	.61
10	3.21	.62	3.28	.70	3.35	.52	3.25	.67	2.79	.60	3.37	.63	3.00	.47	3.14	.69
α	.84		.88		.81		.87		.85		.82		.81		.88	

Table 2. Item means and corrected item-total correlations for each nation (*n* in parentheses) (continued).

	Greece (100)		Hong Kong (1067)		Hungary (158)		India (398)		Indonesia (536)		Iran (802)		Italy (144)		Japan (430)	
Item	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)
1	3.34	.54	2.76	.45	2.98	.61	3.46	.29	3.49	.37	3.36	.46	3.19	.48	2.45	.60
2	2.91	.27	2.35	.41	2.72	.58	2.97	.38	3.67	.25	3.03	.45	2.91	.37	2.13	.57
3	3.17	.24	1.78	.54	2.98	.55	2.88	.41	2.11	.33	2.97	.54	2.99	.35	1.64	.59
4	3.16	.48	2.09	.65	2.64	.61	2.74	.47	2.83	.51	2.80	.57	3.11	.51	1.86	.71
5	2.98	.51	2.03	.64	2.48	.65	2.68	.50	2.94	.55	2.80	.61	3.06	.61	1.73	.69
6	3.39	.49	2.82	.51	3.14	.61	3.54	.42	3.00	.52	3.47	.45	2.40	.50	2.11	.66
7	2.82	.53	2.40	.62	2.90	.61	3.04	.36	2.75	.54	2.96	.54	3.08	.56	1.80	.73
8	2.96	.51	2.35	.52	3.02	.46	3.05	.36	2.95	.60	2.96	.56	3.04	.38	2.16	.74
9	3.20	.42	2.53	.60	2.86	.70	3.25	.39	2.84	.55	3.08	.61	3.22	.42	2.19	.77
10	2.84	.56	1.94	.61	2.88	.73	2.98	.55	3.54	.43	2.78	.58	3.01	.54	2.16	.71
α	.78		.85		.88		.75		.79		.84		.79		.91	

	Korea (147)		Netherlands (911)		Peru (994)		Poland (690)		Portugal (544)		Russia (495)		Spain (399)	
Item	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)	Mean	*r* (it)
1	2.86	.64	3.16	.45	3.02	.33	2.80	.47	3.38	.38	3.38	.58	3.30	.59
2	2.60	.57	2.62	.38	3.61	.36	2.51	.47	3.05	.29	3.23	.56	2.61	.38
3	2.59	.60	3.05	.48	3.20	.33	2.65	.51	2.85	.44	3.09	.60	2.96	.35
4	2.80	.59	3.25	.55	3.11	.54	2.65	.52	2.76	.53	2.85	.52	3.33	.61
5	2.88	.70	3.10	.60	3.16	.55	2.76	.63	2.92	.53	3.31	.65	3.17	.64
6	2.90	.53	3.31	.63	2.97	.57	3.08	.58	3.46	.42	3.40	.60	3.42	.66
7	2.68	.53	3.16	.62	3.10	.59	2.76	.57	2.89	.42	3.06	.52	3.22	.58
8	2.82	.64	3.02	.59	3.43	.45	2.82	.56	2.80	.41	3.38	.36	3.11	.57
9	2.78	.65	3.21	.60	3.15	.50	2.72	.63	3.26	.39	3.11	.56	3.18	.52
10	2.94	.55	3.25	.59	3.17	.47	2.78	.61	2.98	.48	3.11	.53	3.01	.54
α	.88		.85		.80		.85		.76		.85		.84	

	Syria (264)		USA (1594)	
Item	Mean	*r* (it)	Mean	*r* (it)
1	3.16	.42	3.28	.54
2	3.04	.47	2.59	.42
3	2.68	.42	2.62	.44
4	2.88	.40	2.96	.68
5	3.02	.61	2.89	.69
6	2.84	.56	3.33	.59
7	3.02	.38	2.84	.60
8	3.04	.56	2.89	.56
9	2.84	.51	3.05	.64
10	2.48	.34	3.04	.69
α	.79		.87	

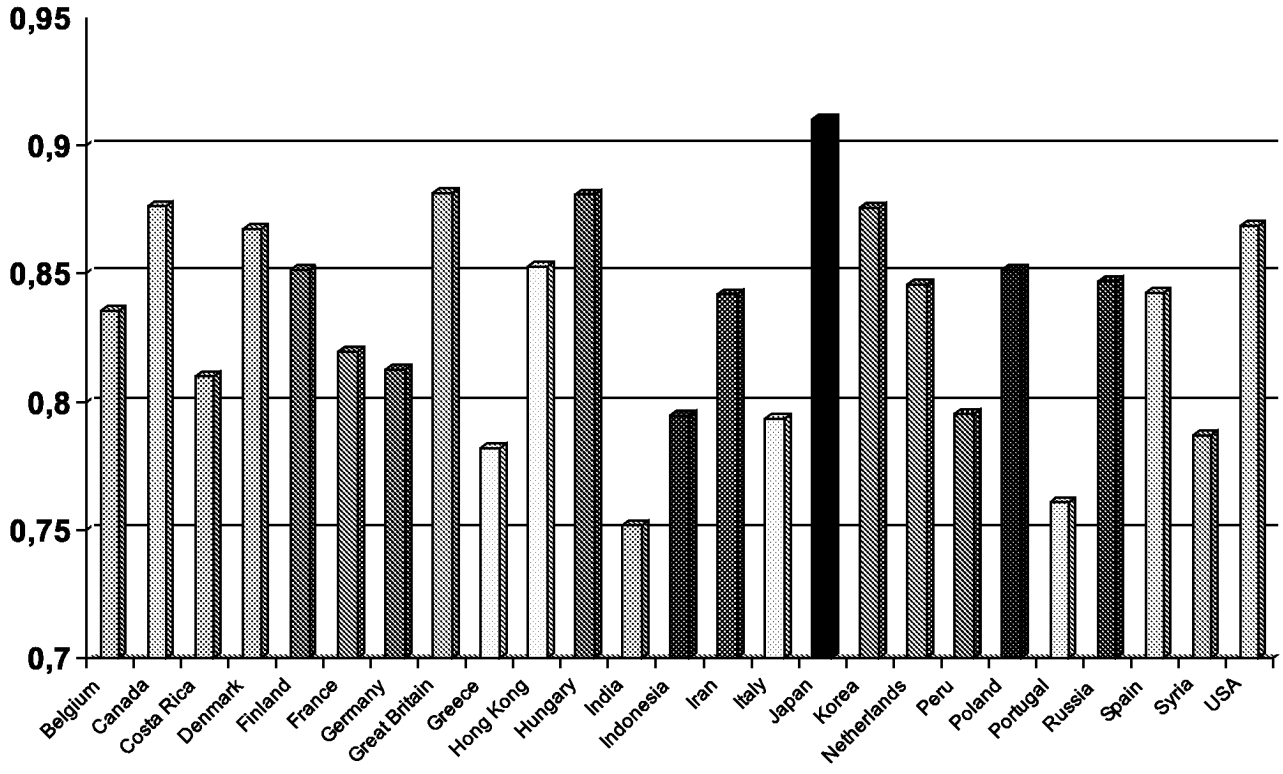

Figure 1. Scale reliabilities (Cronbach's α) for all nations.

Chinese, Indonesian, Japanese, and Korean versions (Zhang & Schwarzer, 1995; Schwarzer, & Born, 1997; Schwarzer, Bäßler, et al., 1997; Schwarzer, Born, et al., 1997; Schwarzer et al., 1999) (all language versions and references are available on line at: www.ralfschwarzer.de/ or www.healthpsych.de).

Sample

The present sample consists of 19,120 persons from 25 countries (7243 men and 9198 women, as well as 2679 participants who did not indicate their sex). The criterion for inclusion in the study was that all ten items had to be answered. Although this is a very conservative criterion, only 3.4% of the original sample did not meet this requirement, which was tolerable. Table 1 displays the number and the mean age of participants, broken down by nation and gender.

The mean age of those 14,634 participants who reported their age was 25 years (SD = 14.7), with a minimum of 12 and a maximum of 94 years. Men (M = 23 years; SD = 11.8; n = 6091) were younger than women (M = 27 years; SD = 16.5; n = 8402). This difference turned out to be significant, $F(1, 14491) = 182{,}99$, $p < .001$, $\eta^2 = .012$, although it accounted for only 1% of the variance. Only 50.4% (n = 9630) of the participants from different countries specified their professions: 34.7% were students, 10% nurses, 3.6% educators, and 2.1% national security service employees, such as police or military.

Results

Item-Level Statistics

Item analyses were carried out separately for each language adaptation. Items had a response range from 1 *(not at all true)* to 4 *(exactly true)*. Item means and corrected item-total correlations are given in Table 2. All item-total correlations, except Item 1 for Costa Ricans and Indians, Item 2 for Greeks and Indonesians, and Item 3 for Belgians and Greeks, turned out to be satisfactory. No overall improvement was possible by the elimination of critical items.

The internal consistency for the total sample (N = 19,120 respondents) was α = .86. Figure 1 displays the reliabilities for all samples. The highest coefficient was found for the Japanese, with α = .91, and the lowest for the Indians, α = .75.

Internal Structure of the Measure

In previous studies, there was strong evidence for assuming that the GSE scale is unidimensional (e. g., Schwarzer & Born, 1997). To confirm the unidimensionality of all scale adaptations, principal component analyses (PCA) were computed separately for each of the 25 subsamples. The Kaiser-Guttman eigenvalue criterion and scree tests indicated one-factor solutions for almost all of the subsamples. Together with the high internal consistencies, these findings suggest unidimensionality within subsamples. Multigroup confirmatory factor analyses were conducted in previous studies that contained only three or four subgroups (e. g., Schwarzer, Bäßler et al., 1997), but this seemed to be inappropriate for the present 25 samples. Therefore, the following analyses involved the total sample.

Confirmatory factor analyses (CFA) were conducted using LISREL Version 8.12a (Jöreskog & Sörbom, 1993a). Initially, principal component analyses were carried out in order to generate both a one-factor and a multifactor solution. The optimal number of factors was determined by combining three criteria:

- The eigenvalues of the factors should be greater than 1.00;
- The multifactor solution should contain at least three loadings for each factor, that is, additional factors with only two or less loadings were unacceptable;
- Each factor should explain at least 5% of the variance.

After that, LISREL maximum likelihood estimation procedures were performed to determine whether a one-factor or a multifactor solution adequately fit the data. Several criteria were used to evaluate the model-data fit: GFI ≥ 0.90 (goodness of fit index, amount of variance and covariance of the model); AGFI > 0.90 (adjusted goodness of fit index with an adjustment in the degrees of freedom); RMR ≤ 0.05 (root mean square residual); RMSEA ≤ 0.08 (root mean square error of approximation); NFI > 0.90 (normed fit index, *F*Symbol"c^2 independence model – *F*Symbol"c^2 specified model/*F*Symbol"c^2 independence model) (Browne & Cudeck, 1993; Jöreskog & Sörbom, 1993b). Because of the large sample size of the present study, *F*Symbol"c^2 – fit statistics are not reported here as they are overly strict and sensitive for models with large samples (Bentler & Bonnett, 1980).

Results of the Kaiser-Meyer-Olkin statistic (KMO = .93) and Bartlett's test of sphericity ($\chi^2 = 56{,}057$, $df = 45$, $p = .00$) indicated that the correlation matrix was suitable for factor analysis. PCA revealed the following loadings for the one-factor solution: .74, .71, .70, .70, .70, .68, .63, .63, .56, .54. In addition, the following eigenvalues were computed: 4.39, .83, .80, .69, etc. For the two-factor solution, PCA gave the following loadings: (a) Factor 1: .74, .71, .70, .70, .70, .69, .63, .63, .56, .54; (b) Factor 2: –.14, –.21, –.15, –.10, .03, .18, –.36, .11, .08, .75.

Thus, the two-factor solution did not meet the criteria. Factor 2 had an eigenvalue of .83, which is below 1.00, and it was constituted by only two loadings of .75 and –.36. Consequently, the multidimension hypothesis was rejected.

The assumption of unidimensionality was supported by a confirmatory factor analysis (CFA). Figure 2 shows the LISREL path diagram as well as factor loadings generated with the maximum likelihood estimation method for the one-factor solution. The following coefficients demonstrate the global goodness of fit: GFI = .98, AGFI = .97, NFI = .97, RMR = .03, and RMSEA = .05. These statistics indicate an excellent fit of the data and the unidimensional model.

Based on these results, it can be concluded that the

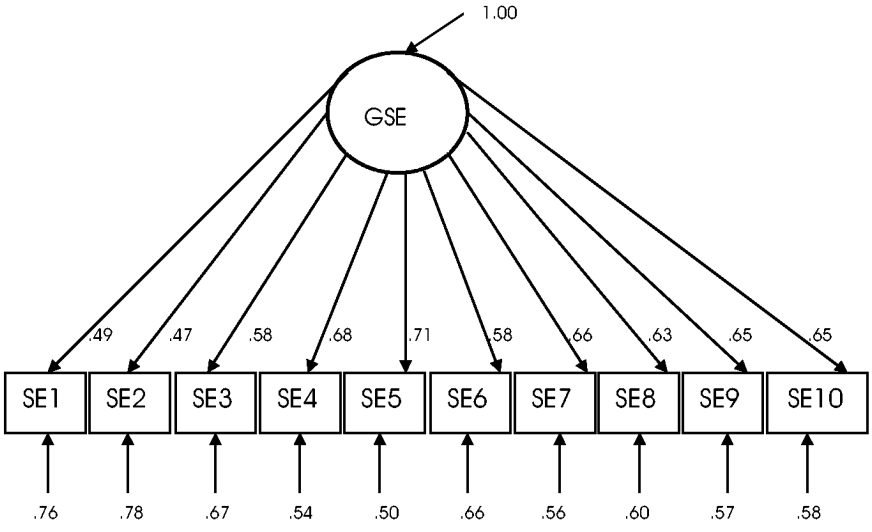

Figure 2. Structural equation model for the test of unidimensionality (GSE = General Self-Efficacy; SE = self-efficacy item).

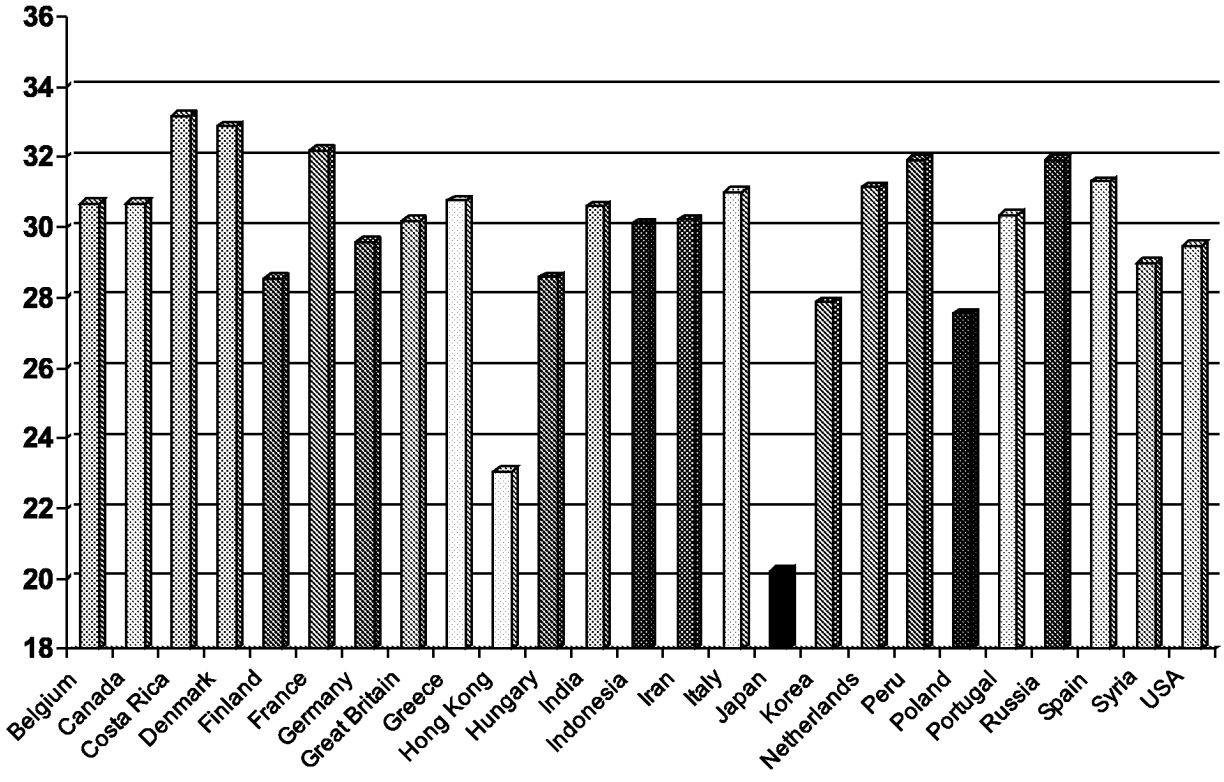

Figure 3. Mean sum scores broken down by nations.

General Self-Efficacy scale is unidimensional and meets the criteria required for multicultural assessment procedures. The results replicate findings from an earlier study based on 13 nations, for which Schwarzer and Born (1997) had also reported a one-factor solution, where eigenvalues of 4.9, .81, .72, and lower were obtained.

Analyses at the Composite Score Level

The possible response range for each item was 1 to 4. Correspondingly, sum scores ranged from 10 to 40. The frequency distribution of the self-efficacy sum scores of the total sample comes close to a normal distribution (M = 29.55, SD = 5.32, kurtosis = .38, skewness = –.52, N = 19,120). This again is in line with previous research (i. e., Schwarzer & Born, 1997). The slight skewness indicates that the scale is more sensitive to detecting individual differences in the lower than in the higher range of the distribution.

Focusing on the mean differences between the international samples, a two-way analysis of variance was computed with nation and gender as factors, whereas self-efficacy composite scores served as the dependent variable. Significant main effects were found for nation, $F(18, 16{,}402) = 238.45$, $p < .001$, $\eta^2 = .21$, and gender, $F(2, 16{,}402) = 24.45$, $p < .001$, $\eta^2 = .003$. An interaction also emerged, $F(18, 16{,}402) = 4.0$, $p < .001$, $\eta^2 = .004$. In light of the low effect sizes of this interaction and the gender main effect, these can be neglected, leaving differences between the nations as the only substantial finding. Figure 3 displays the mean sum scores broken down by nations. The lowest means were found for the Japanese (20.22), followed by the Hong Kong Chinese (23.05). Highest values were found for the Costa Ricans (33.19), Danes (32.87), and French (32.19).

Another two-way analysis of variance was computed, with country and profession (students, nurses, educators and employees of national security services) as the independent variables and self-efficacy sum scores as the outcome measure. The two main effects as well as the interaction turned out to be significant, nation: $F(14, 9{,}611) = 168.13$, $p < .001$, $\eta^2 = .19$, profession: $F(3, 9{,}611) = 19.46$, $p < .001$, $\eta^2 = .006$, nation by profession: $F(1, 9{,}611) = 7.12$, $p < .01$, $\eta^2 = .001$. In terms of explained variance (0.6%), however, the participants' professions did not make a difference in the mean levels of self-efficacy.

Further Evidence of Validity

Evidence of the validity of the GSE scale has been previously published (e. g., Schwarzer & Born, 1997;

Schwarzer et al., 1999). In the present sample, there was a correlation of $r = .07$ with age ($n = 13{,}098$), which is in line with the theory that postulates age-independence of the construct. Further, there are some interesting findings within some of the subsamples. In Costa Rica, for example, GSE correlates $r = -.43$ with anxiety for women ($n = 393$) and $r = -.42$ for men ($n = 258$), with depression $r = -.46$ (women) and $r = -.33$ (men), with optimism $r = .60$ (women) and $r = .52$ (men), and with expected social support $r = .43$ (women) and $r = .30$ (men). In the German subsample, numerous findings corroborated the validity of the instrument (reported above in the introduction).

Discussion

The main goals of the present study were to examine the psychometric properties of the various adaptations of the General Self-Efficacy (GSE) scale and to explore whether self-efficacy can be regarded as a universal construct. When looking for universality of a construct across countries, several prerequisites were considered. First, a culturally sensitive adaptation of the instrument, which is superior to a literal translation, was made for all languages presented here. Second, it had the same structure across cultures, which is especially important when dealing with multidimensional instruments. The minimum condition for this is the "configural equivalence," which means that there should be identical patterns of factors and items across cultures. A more conservative condition that could be added when the configural equivalence is given is "metrical equivalence," which is fulfilled when all items have about the same loadings across samples. Since this is hard to achieve in practice, it is usually seen as satisfactory when the requirement of configural equivalence is met. The present instrument has been designed as a unidimensional measure. Thus, the test of configural equivalence across samples is reduced to the question of whether one can consider empirical unidimensionality within each sample.

Following these considerations, the present research supports the assumption that general perceived self-efficacy is a unidimensional and universal construct. The evidence of its universality originates from the assessment of GSE in 25 countries. Furthermore, these analyses replicate previous results by Schwarzer and Born (1997), who studied the psychometric properties of the general self-efficacy scale with samples from 13 nations. Internal consistencies, item-total correlations, factor loadings, and fit indices of the confirmatory factor analysis indicate that the GSE scale is reliable, homogeneous, and unidimensional across 25 nations.

However, some questions remain open, such as the differences in the GSE sum scores between countries and gender. For example, why do Costa Ricans exhibit the highest GSE level and Japanese the lowest? Or, in general, what are the reasons for the differences in the GSE sum scores between nations? Several explanations are possible: First, these differences actually do exist, and thus the observed estimates are considered to be valid. Results supporting this hypothesis stem from cross-cultural self-efficacy studies, showing that Asian students (i. e., from collectivistic cultures) reported lower self-efficacy beliefs than their non-Asian peers, even though they were more successful in terms of academic achievement (Stigler et al., 1985; Yan & Gaier, 1994). One possible explanation could be that hard work and effort is more highly valued than ability in collectivistic cultures. Therefore, self-efficacy may be rated lower in collectivistic cultures than in individualistic cultures.

Second, the differences in the GSE sum scores between nations could reflect method differences between studies, for example, when the conditions of data collection differed in an uncontrollable manner. Third, they could be due to a population bias since no random samples were drawn from each country. This is one of the most frequent and most problematic shortcomings in cross-cultural studies.

Another question concerns gender differences. In some nations, the GSE levels of women were slightly lower than those of men, but, on the other hand, the interaction between nation and gender documents that this effect is unsystematic. Here again, several explanations seem possible. First, in some cultures there may be genuine GSE differences between men and women, for example due to culturally defined gender roles, and the present results seem to reflect such differences. But, again, unknown methodological problems or nonrepresentative samples may be responsible for the findings.

Developing psychometric tools for different cultures involves a never-ending validation process. If the function of a measure is identical across cultures and within each culture, it would be useless for detecting cultural differences. If, on the other hand, cultural or national differences are found, one cannot be sure whether the instrument's characteristics are responsible for these differences. The present approach was chosen in order to develop a scale that is internally consistent and reliable across nations and languages, which implies configural equivalence (here: unidimensionality). This has been achieved. Thus, it seems likely that mean differences between samples bear cultural importance. What actually constitutes the cultural variance, and to what degree method variance has an impact, remain to be examined in further studies.

Acknowledgments

This paper was an international collaboration that included numerous investigators. The researchers mentioned in the list below either made or supervised the scale adaptation and/or the data collection. We appreciate their continued support and enthusiasm for this project.

Issa Al-Manssour, Syria
Maria Alcalde, Peru
Judith Bäßler, Germany (Spanish version)
Julie Barlow, UK
Aristi Born, Germany (Indonesian version)
Sandra Burigo, Brazil
Susana Coimbra, Portugal
Parvin Daeipour, Iran
Michelle Dumont, Canada
Eva Gemzøe Mikkelsen, Denmark
Ekaterini Glynou, Greece
Esther Greenglass, Canada
Eva Heftyová, Slovakia
Saburo Iwawaki, Japan
Samvel Jeshmaridian, Armenia
Matthias Jerusalem, Germany
Zygfryd Juczyński, Poland
Maria Kopp, Hungary
Marißn Košč, Slovakia
Marcus Koskinen-Hagman, Sweden
Jaro Křivohlavý, Czech Republic
Young-Min Lee, Korea
Lily Moghaddam, Iran
Florian Müller, Germany
John Mueller, Canada
Elahe Nezami, USA (Farsi version)
Renato Nunes, Portugal
Halliki Rimm, Estonia
Espen Røysamb, Norway
Vladimir Romek, Russia
Wilmar Schaufeli, Netherlands
Christine Schwarzer, Germany
Lucio Sibilia, Italy
Ilijana Stamova, Germany
Catherine Tang, China
Bart Teeuw, Netherlands
Shaké Toukmanian, Canada
Kazimierz Wrzesniewski, Poland
Xin Yue, China
Aysin Yesilay, Germany (Turkish version)
Moshe Zeidner, Israel
Jian Xin Zhang, China

References

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

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

Bandura, A. (2000). Exercise of human agency through collective efficacy. *Current Directions of Psychological Science, 9,* 75–78.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52,* 1–26.

Bentler, P.M., & Bonnett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588–606.

Browne, M.W., & Cudeck, R. (1993). Alternatives ways of assessing model fit. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Jöreskog, K.G., & Sörbom, D. (1993a). *Windows LISREL 8.12a.* Chicago: Scientific Software International.

Jöreskog, K.G., & Sörbom, D. (1993b). Testing structural equation modeling. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 294–317). Newbury Park, CA: Sage.

Schröder, K.E.E., Schwarzer, R., & Konertz, W. (1998). Coping as a mediator in recovery from cardiac surgery. *Psychology & Health, 13,* 83–97.

Schwarzer, R. (Ed.). (1992). *Self-efficacy: Thought control of action.* Washington, DC: Hemisphere.

Schwarzer, R., Bäßler, J., Kwiatek, P., Schröder, K., & Zhang, J.X. (1997). The assessment of optimistic self-beliefs: Comparison of the German, Spanish, and Chinese versions of the General Self-Efficacy scale. *Applied Psychology: An International Review, 46,* 69–88.

Schwarzer, R., & Born, A. (1997). Optimistic self-beliefs: Assessment of general perceived self-efficacy in thirteen cultures. *World Psychology, 3,* 177–190.

Schwarzer, R., Born, A., Iwawaki, S., Lee, Y.-M., Saito, E., & Yue, X. (1997). The assessment of optimistic self-beliefs: Comparison of the Chinese, Indonesian, Japanese and Korean versions of the General Self-Efficacy scale. *Psychologia: An International Journal of Psychology in the Orient, 40,* 1–13.

Schwarzer, R., Hahn, A., & Jerusalem, M. (1993). Negative affect in East German migrants: Longitudinal effects of unemployment and social support. *Anxiety, Stress, and Coping, 6,* 57–69.

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35–37). Windsor, UK: NFER-NELSON.

Schwarzer, R., & Jerusalem, M. (Eds.). (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen* [Scales for the assessment of teacher and student characteristics]. Berlin, Germany: Freie Universität Berlin.

Schwarzer, R., Mueller, J., & Greenglass, E. (1999). Assessment of perceived general self-efficacy on the Internet: Data collection in cyberspace. *Anxiety, Stress, and Coping, 12,* 145–161.

Schwarzer, R., Schmitz, G.S., & Tang, C. (2000). Teacher burnout in Hong Kong and Germany: A cross-cultural validation of the

Maslach Burnout Inventory. *Anxiety, Stress, and Coping, 13,* 309–326.

Sherer, M., & Maddux, J.E. (1982). The Self-Efficacy Scale: Construction and validation. *Psychological Reports, 51,* 663–671.

Skinner, E.A., Chapman, M., & Baltes, P.B. (1988). Control, means-ends, and agency beliefs: A new conceptualization and its measurement during childhood. *Journal of Personality and Social Psychology, 54,* 117–133.

Stigler, J.W., Smith, S., & Mao, L.-W. (1985). The self-perception of competence by Chinese children. *Child development, 56,* 1259–1270.

Yan, W., & Gaier, E.L. (1994). Causal attributions for college success and failure: An Asian-American comparison. *Journal of Cross-Cultural Psychology, 25,* 146–158.

Zhang, J.X., & Schwarzer, R. (1995). Measuring optimistic self-beliefs: A Chinese adaptation of the General Self-Efficacy scale. *Psychologia, 38,* 174–181.

Ralf Schwarzer
Gesundheitspsychologie
Freie Universität Berlin
Habelschwerdter Allee 45
D-14195 Berlin
Germany
Fax +49 30 8385-5634
E-mail health@zedat.fu-berlin.de

Appendix

The General Perceived Self-Efficacy Scale

1 I can always manage to solve difficult problems if I try hard enough.
2 If someone opposes me, I can find the means and ways to get what I want.
3 I am certain that I can accomplish my goals.
4 I am confident that I could deal efficiently with unexpected events.
5 Thanks to my resourcefulness, I can handle unforeseen situations.
6 I can solve most problems if I invest the necessary effort.
7 I can remain calm when facing difficulties because I can rely on my coping abilities.
8 When I am confronted with a problem, I can find several solutions.
9 If I am in trouble, I can think of a good solution.
10 I can handle whatever comes my way.

Response Format:

1 = Not at all true 2 = Hardly true 3 = Moderately true 4 = Exactly true

Note: The English version was developed in 1985, published in 1995, and revised slightly in 2000 (Schwarzer & Jerusalem, 1995).