Research Article

Psychometric Evidence of a Measure of General Self-Efficacy in Peruvian Schoolchildren: Internal Structure, Reliability and Factorial Invariance and Proposal of a Brief Version

Walter L. Arias Gallegos\textsuperscript{a}, Tomás Caycho-Rodríguez\textsuperscript{b}, Lindsey W. Vilca\textsuperscript{c}, Ana Cecilia Quispe Álvarez\textsuperscript{d}, Adolfo Alexander Gamero Díaz\textsuperscript{d}, Rosario Irma Butrón Ortiz\textsuperscript{d}, Mariel Delgado-Campusano\textsuperscript{e}, Mario Reyes-Bossio\textsuperscript{e}

\textsuperscript{a} School of Psychology, Universidad Católica San Pablo, Arequipa, Perú.
\textsuperscript{b} Faculty of Psychology, Universidad Científica del Sur, Lima, Perú.
\textsuperscript{c} South American Center for Education and Research in Public Health, Universidad Norbert Wiener, Lima, Perú.
\textsuperscript{d} National University of San Agustín, Arequipa, Perú.
\textsuperscript{e} Faculty of Psychology, Universidad Peruana de Ciencias Aplicadas, Lima, Peru.

Abstract

The present research aims to propose a brief version of a measure of general self-efficacy, based on the Self-Efficacy Scale, and to evaluate its evidence of validity based on internal structure, reliability and factorial invariance according to sex. A total of 806 students (67.5\% male and 32.5\% female) of secondary level from 13 educational institutions in the city of Arequipa (Peru) participated. The data were processed using Confirmatory Factor Analysis and Item Response Theory. The results indicate that the original 20-item version of the 10-item Brief Alegre Self-Efficacy Scale is unidimensional with adequate goodness-of-fit and reliability indices ($\omega = .89$). In addition, the presence of factorial invariance between males and females was tested. It is concluded that the proposed measure is valid and reliable, with a monotonic level of difficulty.
However, more research is needed with representative samples and with respect to other educational variables.

*Keywords*: self-efficacy; psychometrics; factorial invariance; item response theory.

**Table of Contents**

Method
Results
Discussion
Conclusion
References

**Psychological Thought, 2024, Vol. 17(1), 58-81, [https://doi.org/10.37708/psyct.v17i1.812](https://doi.org/10.37708/psyct.v17i1.812)**

Received: 2023-02-01. Accepted: 2024-01-25. Published (VoR): 2024-04-30.
Handling Editor: Marius Drugaș, University of Oradea, Romania. *Corresponding author at: Universidad Científica del Sur, Lima, Perú. E-mail: tcaycho@cientifica.edu.pe

This is an open access article distributed under the terms of the Creative Common Attribution License ([https://creativecommons.org/licenses/by/4.0](https://creativecommons.org/licenses/by/4.0)), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Self-efficacy is a recent construct that was introduced to the psychological literature by Albert Bandura (1925-2021) after having developed his theory of vicarious learning ([Bandura, 1965](#)), which is based, at first, on the principles of behavior modification ([Bandura, 1983b](#)), to later take a social cognitive orientation ([Bandura, 1969](#), [1982](#); [Bandura & Walters, 1974](#)). However, in later developments, he manages to integrate vicarious experiences as sources that guide self-efficacy beliefs ([Bandura, 1976](#), [2009](#)).

Self-efficacy, then, as conceived by Bandura ([1997](#)) is the set of judgments that people have about their abilities to achieve a certain level of performance, but it encompasses cognitive, affective, motivational, social and psychophysiological processes ([Bandura, 2009](#)). In this sense, as stated by Bandura ([1977](#)), self-efficacy unifies various behavioral, cognitive and social psychological constructs such as reinforcement, expectations, motivation, locus of control, goals, emotional arousal, vicarious experiences, verbal persuasion, coping, social comparison, etc. ([Bandura, 1983a](#), [1986](#), [1987](#)).
While self-efficacy can be considered a belief that people have about their efficacy in general, it is more common for people to demonstrate efficacy in certain tasks or contexts more than in others (Bandura, 2006). Thus, self-efficacy is applied in different facets of human endeavors, such as health (e.g., Isaac et al., 2018; Jiang et al., 2019; Zhou et al., 2021), work, and business (e.g., Srimulyani, & Hermanto, 2021; Gangloff & Mazilescu, 2017; Robalino & Musso, 2022), education (e.g., Hyytinen et al., 2018; Manzano-Sanchez et al., 2018), sports (e.g., Brace et al., 2020; López-Aguayo & Reyes-Bossio, 2018), and many other domains such as family and collectivity (Dimitropoulos et al., 2018; Olatonji et al., 2020).

In the educational context, several studies have shown that self-efficacy is positively related to academic goals and achievement motivation (Valle et al., 2015), reflective learning styles (Laffita & Guerrero, 2017), metacognition (Taghani & Razavi, 2021), productive coping styles (Piergiovanni & Depaula, 2018), flow and engagement (Mesurado et al., 2015), self-concept and psychological well-being (Veliz-Burgos & Apodaca, 2012). While, on the other hand, it is negatively related to academic stress (Jersusalem & Mittag, 2009), substance addictions in adolescents (Marlatt et al., 2009), school failure (Lata & Nasa, 2014), anxiety, neuroticism, somatization, depression, hostility, phobias, paranoid thinking and psychoticism (García-Méndez, & Rivera-Ledesma, 2020), among other variables related to mental health.

In this regard, there are several studies that highlight various explanatory factors of self-efficacy, depending on the gender of students and their cognitive abilities (Vasile et al., 2011), the nationality of students, mediated by their sociocultural differences (Steca et al., 2009; Mesurado et al., 2015); personality traits (García-Méndez, & Rivera-Ledesma, 2020; Stajkovic et al., 2018) and self-concept (Merino & Díaz, 2003). In this regard, Bandura (1993) explained that students' beliefs in their own abilities have effects on their academic performance and vocational decisions, since self-efficacy is a determinant in the processes of self-regulation of learning and the use of metacognitive strategies, necessary for the academic orientation of cognitive functioning. Given that self-efficacy has an impact on emotional regulation, it also affects the school socialization processes inherent to interactions among peers and with teachers, as well as the promotion of an adequate learning climate (Bandura, 2009).

Consequently, the evaluation of self-efficacy becomes fundamental in educational contexts. In Peru, however, there are still few studies on self-efficacy, although it has been shown that self-efficacy is positively related to self-regulation of learning and academic performance (Alegre,
2014) and negatively related to academic procrastination (Alegre, 2013). Other studies have reported that the academic self-efficacy of university students in Peru, reaches a low level in 17.4% of the cases, moderate in 61.2% and a high level in 21.4% (Dominguez-Lara & Fernández-Arata, 2019). More recently, a study with 582 Peruvian university students between 16 and 41 years old reported that males had higher levels of psychological distress, anxiety and academic self-efficacy than females, but in addition, academic self-efficacy predicted satisfaction with their studies (Carranza et al., 2022).

On the other hand, psychometric studies have also been conducted on various self-efficacy scales that are applied in educational contexts or are related to academic variables. Dominguez-Lara (2017) for example, constructed a self-efficacy scale for research. The same author has also reported the psychometric properties of Palenzuela’s Academic Situations Specific Perceived Self-Efficacy Scale (EAPESA) (Dominguez-Lara, 2014; Dominguez-Lara et al., 2012). In a recent study with this scale, evidence has been found that supports the convergence of the scores with those of the Brief Scale Study Satisfaction, while data was also obtained on the divergence between the EAPESA score and psychological distress (Carranza et al., 2022).

Another scale that has been used in Peru, whose psychometric properties are reported, is the General Academic Self-Efficacy Questionnaire, which was constructed by Torre (2006), and validated by Alegre (.). A test that has been created in Peru that assesses teacher self-efficacy was constructed by the Educational Quality Measurement Unit of the Peruvian Ministry of Education and validated by Burga (.). The test consists of 12 items with a Likert-type scale with seven response levels, which is applied individually or collectively.

Another scale that has been created and validated in Peru is the Alegre Self-Efficacy Scale (), which was applied to 348 university students in Lima and consists of 20 items, distributed on a Likert-type response scale with five levels. This test was validated by exploratory factor analysis, using the principal components method, reporting a single factor that explains 50.56% of the total variance of the test. Reliability was also calculated using the internal consistency method, and Cronbach’s alpha test, which obtained high scores ($\alpha = .948$). It was also reported, that there is evidence of divergent validity, since the score of this scale correlated negatively with the scores of the Academic Procrastination Scale ($r = -.234$) (Alegre, 2013).

In a psychometric investigation of the Alegre Self-Efficacy Scale with a sample of university students from the city of Arequipa, it was reported that it presents a unidimensional structure that
explains 48.7% of the total variance, by means of the exploratory factor analysis practiced with the robust unweighted least squares (RULS) extraction method. It also presented a high reliability index ($\alpha = .925$) and adequate goodness-of-fit indices ($\chi^2/gl = .973$; $CFI = 1.001$; $GFI = 0.984$; $AGFI= 0.982$; $RMSEA= .000$), calculated using the FACTOR program (Arias & Rivera, 2018).

Having said all this, it can be seen that although there are some studies on the variable self-efficacy in Peru, and several of these are psychometric in nature, in most cases, they have worked based on the classical theory of tests, as well as with samples of university students; leaving aside samples of school students, in whom self-efficacy is also a relevant variable for their emotional self-regulation (Bandura, 2001; Bandura et al., 2003), which predicts their academic performance and cognitive functioning (Bandura, 2009), as well as with university students (Kostagiolas et al., 2019). Consequently, the present instrumental research (Ato et al., 2013), aims to propose a brief version of a measure of general self-efficacy, based on Alegre's (2013) Self-Efficacy Scale, and to evaluate its evidence of validity based on internal structure, reliability and factorial invariance according to gender in a sample of high school students from the city of Arequipa.

## Method

### Participants

The sample consisted of 806 adolescents between 12 and 18 years of age from 13 secondary schools in the city of Arequipa. In addition, 56.6% were between 12 and 14 years old, while 43.4% were between 15 and 18 years old. As for their level of studies, 27.4% are in the first year of secondary school, 22.5% are in the second year, 15.5% in the third year, 13.8% in the fourth year and 20.8% in the fifth year of secondary school. The sample was selected non-probabilistically using the intact group technique (Hernández et al., 2010).

### Instrument

The Self-Efficacy Scale, designed by Alegre (2013), was applied, consisting of 20 items with five response alternatives on a Likert scale from "Totally disagree" (0) to "Totally agree" (4). In a previous study, with a sample of university students from Arequipa, it has been confirmed that the scale has a single factor structure that explains 48.7% of the total variance of the test and has a reliability index of .925, estimated by Cronbach's alpha test (Arias & Rivera, 2018).

### Procedure
The instrument was applied virtually by means of an answer template in Google forms, after coordination with the authorities of the educational institutions and once the parents were informed of the execution of the information and gave their consent for the participation of their children. For this purpose, the confidentiality of their data was guaranteed and the objectives of the study were explained, so that the students who formed the total sample were not forced to participate and had the consent of their parents.

Data Analysis

To evaluate the validity based on the internal structure, the Confirmatory Factor Analysis (CFA) was used, for which the estimator *Diagonally Weighted Least Squares with Mean and Variance corrected* (WLSMV) since the items were ordinal in nature (Brown, 2015). Regarding the fit indices, the chi-square test ($\chi^2$), the RMSEA index and the SRMR index were used, where values lower than .05 indicate good fit, and between .05 and .08 is considered acceptable (Kline, 2015). Also, the CFI and TLI index were used, where values greater than .95 evidenced good fit and greater than .90 an acceptable fit (Jordan Muiños, 2021; Schumacker & Lomax, 2015). Regarding the internal consistency of the scale, Cronbach's alpha coefficient (Cronbach, 1951) and the omega coefficient (McDonald, 1999) were used, where a value of $\omega > .80$ is adequate (Raykov & Hancock, 2005).

To evaluate the factorial invariance of the scale according to the sex of the participants, a sequence of hierarchical variance models was used. First, configural invariance (reference model) was evaluated, followed by metric invariance (equality of factor loadings), scalar invariance (equality of factor loadings and thresholds) and finally strict invariance (equality of factor loadings, thresholds and residuals). To compare the sequence of models we first employed a formal statistical test, for which the chi-square difference ($\Delta\chi^2$) was used, where non-significant values ($p > .05$) suggest invariance between groups. Secondly, a modeling strategy was employed, for which the differences in RMSEA ($\Delta$RMSEA) was used, where differences less than $< .015$ evidence model invariance between the groups (Chen, 2007).

For Item Response Theory (IRT), a Graded Response Model (GRM, Samejima, 1997) was employed specifically an extension of the 2-parameter logistic model (2-PLM) for ordered polytomous items (Hambleton et al., 2010). The C2 test developed for ordinal items (Cai & Monroe, 2014) was used to estimate the model fit and the following fit criteria were used: RMSEA $\leq .05$ (Maydeu-Olivares & Joe, 2014) and SRMSR $\leq .05$ (Maydeu-Olivares, 2013). CFI and TLI
values were also taken into account using the same fit criterion ($\geq .95$) employed in SEM models (Lubbe & Schuster, 2019).

Two types of parameters were estimated for each item, discrimination (a) and difficulty (b). The discrimination parameter (a) determines the slope at which item responses change as a function of the level in the latent trait and the item difficulty parameters (b) determine how much of the latent trait the item requires to be answered. Since the scale has five response categories, there are four difficulty estimates, one per threshold. The estimates for these four thresholds indicate the level of the latent variable at which an individual has a 50% chance of scoring at or above a particular response category. Information Curves were also calculated for the items and the scale.

All statistical analyses were performed using the "lavaan" package (Rosseel, 2012) for the AFC, the "semTools" package (Jorgensen et al., 2018) for factorial invariance and the "mirt" package for the GRM (Chalmers, 2012). In all cases, the RStudio environment (RStudio Team, 2018) was used for R (R Core Team, 2019).

**Results**

**Descriptive Analysis**

Table 1 shows that item 6 ("I can solve most problems if I try hard enough") has the highest mean score in the sample ($M = 4.20$). It can also be seen that item 10 ("No matter what I have to face, I am usually prepared for it") presents the lowest mean score in the sample ($M = 3.67$). With respect to the skewness and kurtosis indices, it can be seen that the items present adequate indices (As < ±2; Ku < ±7), in accordance with the criteria of Finney and DiStefano (2013).
<table>
<thead>
<tr>
<th>Ítems</th>
<th>Descriptive Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
</tr>
<tr>
<td>1. En general, puedo manejar las dificultades si me esfuerzo.</td>
<td>4.17</td>
</tr>
<tr>
<td>2. Cuando alguien interfiere con lo que quiero, puedo encontrar alternativas para conseguir lo que busco.</td>
<td>4.03</td>
</tr>
<tr>
<td>3. Se me hace fácil mantener mis objetivos para así lograr mis metas.</td>
<td>3.75</td>
</tr>
<tr>
<td>4. Tengo confianza en mí mismo de poder enfrentarme a situaciones inesperadas.</td>
<td>3.90</td>
</tr>
<tr>
<td>5. Gracias a mis cualidades, sé cómo manejar situaciones imprevistas.</td>
<td>3.87</td>
</tr>
<tr>
<td>6. Puedo resolver la mayoría de los problemas si me esfuerzo lo suficiente.</td>
<td>4.20</td>
</tr>
<tr>
<td>7. Consigo mantener la calma cuando enfrento dificultades porque confío en mis habilidades.</td>
<td>3.78</td>
</tr>
<tr>
<td>8. Cuando enfrento un problema, soy capaz de encontrar varias soluciones.</td>
<td>3.84</td>
</tr>
<tr>
<td>9. Si estoy en un aprieto, usualmente pienso alternativas para resolver la situación.</td>
<td>4.05</td>
</tr>
<tr>
<td>10. Sin importar lo que tenga que enfrentar, por lo general estoy preparado para ello.</td>
<td>3.67</td>
</tr>
<tr>
<td>11. Puedo encontrar la manera de obtener lo que quiero, aunque alguien se oponga.</td>
<td>3.75</td>
</tr>
<tr>
<td>12. Sucedan lo que sucedan, me considero capaz de manejar la situación.</td>
<td>3.79</td>
</tr>
<tr>
<td>13. Cuando me encuentro en una situación difícil, genero alternativas de acción.</td>
<td>3.84</td>
</tr>
<tr>
<td>14. Gracias a mis recursos, puedo superar situaciones espontáneas.</td>
<td>3.77</td>
</tr>
<tr>
<td>15. Me planteo objetivos en los que persisto para alcanzar mis metas.</td>
<td>3.99</td>
</tr>
<tr>
<td>16. Cuando me presentan problemas mantengo la calma pues confío.</td>
<td>3.81</td>
</tr>
<tr>
<td>17. Por más difícil que parezca algún problema que se me presente, pienso en algunas conductas para afrontarlo óptimamente.</td>
<td>3.88</td>
</tr>
<tr>
<td>18. Si algo se opone en el logro de mis metas, planteo alternativas de solución.</td>
<td>4.05</td>
</tr>
<tr>
<td>19. Puedo generar diversas respuestas ante los problemas que se me presentan cotidianamente.</td>
<td>3.92</td>
</tr>
<tr>
<td>20. El asumir como reto los problemas que se me presentan me lleva a movilizar mis recursos para solucionarlos.</td>
<td>3.93</td>
</tr>
</tbody>
</table>
Validity based on internal structure and reliability

The unidimensional model of twenty items was found to present adequate fit indices in the total sample of participants $\chi^2(170) = 674.01, \ p < .001; \ RMSEA = .061 \ [CI90\% .056 - .066]; \ SRMR = .037; \ CFI = .96; \ TLI = .96)$. However, Table 2 shows that there are items with a moderate factorial weight (.52 to .69) and with a high level of measurement error (.52 to .73). Therefore, items with a factorial weight $> .70$ and with a measurement error $< .49$ were selected to develop a brief version of the Self-Efficacy Scale.

The new ten-item version presented adequate fit indices in the total sample of participants $\chi^2(35) = 158.82, \ p < .001; \ RMSEA = .066 \ [CI90\% .056 - .077]; \ SRMR = .030; \ CFI = .98, \ TLI = .98$. It can also be seen in Table 3 that the brief version evidences acceptable fit indices in the specific group: women $\chi^2(35) = 93.42, \ p < .001; \ RMSEA = .080 \ [CI90\% .061 - .100]; \ SRMR = .039; \ CFI = .98; \ TLI = .98$ and males $\chi^2(35) = 48.80, \ p < .001, \ RMSEA = .067 \ [CI90\% .000 - .044]; \ SRMR = .028; \ CFI = .98; \ TLI = .98$.

Table 2.
Factorial weights of the full and abbreviated version

<table>
<thead>
<tr>
<th>Items</th>
<th>Complete Version</th>
<th>Abbreviated Version</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda$ (error)</td>
<td>$\lambda$ (error)</td>
</tr>
<tr>
<td>1</td>
<td>.59 (.66)</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>.56 (.69)</td>
<td>–</td>
</tr>
<tr>
<td>3</td>
<td>.58 (.66)</td>
<td>–</td>
</tr>
<tr>
<td>4</td>
<td>.73 (.47)</td>
<td>.73 (.48)</td>
</tr>
<tr>
<td>5</td>
<td>.73 (.46)</td>
<td>.74 (.46)</td>
</tr>
<tr>
<td>6</td>
<td>.60 (.64)</td>
<td>–</td>
</tr>
<tr>
<td>7</td>
<td>.69 (.52)</td>
<td>–</td>
</tr>
<tr>
<td>8</td>
<td>.74 (.45)</td>
<td>.74 (.45)</td>
</tr>
<tr>
<td>9</td>
<td>.67 (.56)</td>
<td>–</td>
</tr>
<tr>
<td>10</td>
<td>.68 (.53)</td>
<td>–</td>
</tr>
<tr>
<td>11</td>
<td>.52 (.73)</td>
<td>–</td>
</tr>
<tr>
<td>12</td>
<td>.76 (.42)</td>
<td>.77 (.40)</td>
</tr>
<tr>
<td>13</td>
<td>.73 (.44)</td>
<td>.72 (.48)</td>
</tr>
<tr>
<td>14</td>
<td>.65 (.58)</td>
<td>–</td>
</tr>
<tr>
<td>15</td>
<td>.68 (.54)</td>
<td>–</td>
</tr>
<tr>
<td>16</td>
<td>.73 (.47)</td>
<td>.73 (.47)</td>
</tr>
<tr>
<td>17</td>
<td>.75 (.43)</td>
<td>.78 (.41)</td>
</tr>
<tr>
<td>18</td>
<td>.73 (.47)</td>
<td>.72 (.48)</td>
</tr>
<tr>
<td>19</td>
<td>.73 (.48)</td>
<td>.73 (.47)</td>
</tr>
<tr>
<td>20</td>
<td>.72 (.48)</td>
<td>.71 (.49)</td>
</tr>
<tr>
<td>Reliability</td>
<td>$\alpha$</td>
<td>.93</td>
</tr>
<tr>
<td></td>
<td>$\omega$</td>
<td>.93</td>
</tr>
</tbody>
</table>

Note. $\lambda =$ factorial weight
Table 2 shows that the brief scale presents adequate reliability indices ($\alpha = .89; \omega = .89$). Similarly occurs in the specific groups: females ($\alpha = .90; \omega = .91$) and males ($\alpha = .88; \omega = .88$).

**Factor invariance by sex**

Table 3 shows that the factor structure of the brief scale has shown evidence of being strictly invariant for the male and female groups in the sequence of invariance models proposed: metric invariance ($\Delta$RMSEA = .006), scalar invariance ($\Delta$RMSEA = .005) and strict invariance ($\Delta$RMSEA = -.003). It is concluded, therefore, that the items of the self-efficacy scale are understood in the same way by males and females.

**Table 3. Invariance models according to sex of participants**

<table>
<thead>
<tr>
<th>One-dimensional model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$p$</th>
<th>SRMR</th>
<th>TLI</th>
<th>CFI</th>
<th>RMSEA</th>
<th>$\Delta\chi^2$</th>
<th>$\Delta$df</th>
<th>$\Delta$p</th>
<th>$\Delta$RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Females</td>
<td>93.42</td>
<td>35</td>
<td>.000</td>
<td>.039</td>
<td>.98</td>
<td>.98</td>
<td>.080</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Males</td>
<td>48.80</td>
<td>35</td>
<td>.061</td>
<td>.028</td>
<td>.98</td>
<td>.98</td>
<td>.027</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Configural</td>
<td>96.27</td>
<td>70</td>
<td>.020</td>
<td>.029</td>
<td>.97</td>
<td>.98</td>
<td>.031</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Metric</td>
<td>121.39</td>
<td>79</td>
<td>.002</td>
<td>.040</td>
<td>.96</td>
<td>.97</td>
<td>.037</td>
<td>16.46</td>
<td>9</td>
<td>.058</td>
<td>.006</td>
</tr>
<tr>
<td>Scalar</td>
<td>150.22</td>
<td>88</td>
<td>.000</td>
<td>.043</td>
<td>.95</td>
<td>.95</td>
<td>.042</td>
<td>24.43</td>
<td>9</td>
<td>.004</td>
<td>.005</td>
</tr>
<tr>
<td>Strict</td>
<td>157.52</td>
<td>98</td>
<td>.000</td>
<td>.047</td>
<td>.96</td>
<td>.95</td>
<td>.039</td>
<td>12.14</td>
<td>10</td>
<td>.276</td>
<td>-.003</td>
</tr>
</tbody>
</table>

**Item Response Theory Model: Graded Response Model (GRM)**

The results found in the Confirmatory Factor Analysis (CFA) allow us to fulfill the two main assumptions: the existence of unidimensionality and consequently local independence. Therefore, a Graded Response Model (GRM) was used, specifically an extension of the 2-parameter logistic model (2-PLM) for ordered polytomous items. Table 4 shows that the GRM model presents adequate fit indices $\chi^2(35) = 104.82, p < .001$; RMSEA = .049; SRMSR = .034; TLI = .99; CFI = .99. It is also appreciated that all item discrimination parameters are above the value of 1, generally considered as good discrimination (Hambleton et al., 2010). Regarding the difficulty parameters, all threshold estimators increased monotonically. That is, a greater presence of the latent trait is required to answer the higher response categories.
Table 4.
GRM model fit indices, discrimination and difficulty parameters for the brief scale items.

<table>
<thead>
<tr>
<th>Item</th>
<th>Items parameters</th>
<th>GRM Model Fit Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a</td>
<td>b₁</td>
</tr>
<tr>
<td>Item 4</td>
<td>1.97</td>
<td>-2.89</td>
</tr>
<tr>
<td>Item 5</td>
<td>2.07</td>
<td>-3.62</td>
</tr>
<tr>
<td>Item 8</td>
<td>2.11</td>
<td>-3.58</td>
</tr>
<tr>
<td>Item 12</td>
<td>2.33</td>
<td>-3.11</td>
</tr>
<tr>
<td>Item 13</td>
<td>1.99</td>
<td>-3.68</td>
</tr>
<tr>
<td>Item 16</td>
<td>1.98</td>
<td>-3.57</td>
</tr>
<tr>
<td>Item 17</td>
<td>2.19</td>
<td>-3.94</td>
</tr>
<tr>
<td>Item 18</td>
<td>1.67</td>
<td>-4.16</td>
</tr>
<tr>
<td>Item 19</td>
<td>2.01</td>
<td>-4.08</td>
</tr>
<tr>
<td>Item 20</td>
<td>1.99</td>
<td>-3.53</td>
</tr>
</tbody>
</table>

Note. a = discrimination parameters; b = difficulty parameters.

Figure 2 shows the Information Curves for the ten items and the scale (IIC and ICT respectively). The IIC shows that items 12 and 17 are the most accurate items of the scale for assessing the latent trait. In addition, the TIC shows that the test is more reliable (accurate) in the range of the scale between -4 and 1.5.
Discussion

Self-efficacy has been the object of study in psychology for more than 40 years since 1977, when Bandura first mentioned this construct (Velásquez, 2012). Numerous studies have proven its usefulness in the prediction of various psychological variables in clinical, occupational, social and educational contexts (Bandura, 2009). In education, self-efficacy is a variable that can be applied to the study of teaching performance (Sánchez-Rosas et al., 2021) as well as to the cognitive functioning of students, since it is linked to behavioral self-regulation, metacognition and students’ vocational decisions (Bandura, 1993).

In Peru, self-efficacy has been recently studied from a preferably psychometric perspective, through the analysis of various scales that measure teaching self-efficacy (Burga, 2015), academic self-efficacy (Alegre, 2014; Domínguez-Lara, 2014; Domínguez-Lara et al., 2012, Domínguez-Lara & Fernández-Arata, 2019), research self-efficacy (Domínguez-Lara, 2017), and self-efficacy in general (Alegre, 2013; Arias & Rivera, 2018). These investigations, however, have been based in the vast majority of cases on samples of university students in the city of Lima (capital of Peru), and have used models derived from Classical Test Theory.
Initially, a unidimensional structure with a high reliability index of Alegre Self-Efficacy Scale was reported, which is consistent with a later study in university students in Arequipa (Arias & Rivera, 2018). However, the initial study by Alegre (2013) and that of Arias and Rivera (2018) are based on classical theory and apply an exploratory factor analysis; whereas, the present study applies confirmatory factor analysis and analyses corresponding to Item Response Theory, which are based on different psychometric principles to assess unidimensionality (Burga, 2006). Based on this, it was observed that items 1, 2, 3, 6, 7, 9, 10, 11, 14 and 15 obtained moderate factorial weights with a high level of measurement error. Faced with this, it was decided to select items with a factorial weight > .70 and with a measurement error < .49 to develop a brief version of the Self-efficacy Scale (Alegre, 2013).

This new version of only 10 items, in addition to presenting better goodness-of-fit indices (Jordan Muiños, 2021), shows acceptable values of factorial invariance between males and females, which is an indicator that the test can be understood and applied to secondary school students, regardless of their sex. Although, on the other hand, the internal consistency indices were higher for the 20-item version, both with Cronbach’s alpha test and McDonald’s Omega test, the abbreviated version also obtained adequate reliability indices. Therefore, its use can be suggested in educational contexts to assess self-efficacy as a valid and reliable measure of cognitive functioning of schoolchildren in the city of Arequipa, and in relation to other associated variables such as metacognitive skills (Bandura, 2009), academic achievement motivation (Arias et al., 2018), learning styles (Laffita & Guerrero, 2017), self-regulated learning (Arias et al., 2020), etc.

Regarding the application of Item Response Theory, confirmatory factor analysis suggests compliance with the assumptions of unidimensionality and local independence. The Graded Response Model (GRM) presented adequate fit indices and all item discrimination parameters are good (Hambleton et al., 2010), while the difficulty parameters increased monotonically. Thus, items 12 and 17 of the abbreviated scale are the most accurate for assessing self-efficacy in general, however, it should be noted that, as suggested by Bandura (2006), it is advisable to assess self-efficacy specifically.

In the case of the educational context, academic self-efficacy, conceived as a set of judgments about one’s own abilities required to organize and execute actions in academic settings
(Domínguez-Lara et al., 2012), turns out to be a more appropriate measure in school-based education than self-efficacy as a general variable. In addition, academic self-efficacy has a multidimensional composition (García-Méndez & Rivera-Ledesma, 2020), since from an agentic conception (Bandura, 2001) several factors such as intentionality, the formulation of academic goals, planning, the use of metacognitive strategies, among others, can be recognized. Likewise, given that school teaching is nowadays carried out, also by virtual e-learning means, it would be advisable, to include aspects related to skills for the use of technologies in self-efficacy (DeNoyeles et al., 2014).

**Limitations and Future Research Suggestions**

The study is not free of limitations. First, the original, brief version of Alegre's (2013) Self-Efficacy Scale does not specifically assess academic self-efficacy. Second, the total sample has not been selected by probabilistic methods, which prevents the generalization of results. Therefore, it would be advisable to have a more representative sample of the Arequipa population for subsequent studies. Nevertheless, our results show the psychometric properties of a self-efficacy scale that could be used as a screening measure or for research purposes in educational contexts.

**Conclusion**

Finally, it is necessary to carry out further research on the psychometric properties of the Self-Efficacy Scale through convergent, divergent, discriminant, etc. validity. In this sense, this work offers the possibility of using the Self-Efficacy Scale and assessing its relationships with other psychological variables within the educational environment at the school level, given that in Peru the psychometric properties of the self-efficacy scales created or validated at the national level have not been analyzed at this level of education. One can conclude that the Alegre Self-Efficacy Scale (2013), is a valid and reliable measure, which does not offer sex-discriminatory measures, recommending its use as a 20-item scale as a short 10-item measure, for the Peruvian context.

**List of Abbreviations**

RMSEA: Root mean square error of approximation  
SRMR: standardized root mean square residual  
CFI: Comparative fit index  
TLI: Tucker-Lewis index
ΔRMSEA: Differences in RMSEA
IRT: Item Response Theory
GRM: Graduated Response Model

**Funding/Financial Support**
The authors have no funding to report.

**Other Support/Acknowledgement**
The authors have no support to report.

**Competing Interests**
The authors have declared that no competing interests exist.
References


Psychometric Evidence of a Measure of General Self-efficacy


Psychometric Evidence of a Measure of General Self-efficacy


Psychometric Evidence of a Measure of General Self-efficacy


**About the authors**

**Walter L. Arias Gallegos**, Doctor of Psychology, San Agustín National University, Arequipa, Peru. Researcher at the San Pablo Catholic University, Arequipa, Peru.

**Tomás Caycho-Rodríguez**, Ph.D. in Psychology. Senior Researcher at the Universidad Científica del Sur, Research Renacyt Distinguished category. His research interests are psychometrics and cross-cultural research.
Lindsey W. Vilca, Master in Psychology. Professor and researcher at the Norbert Wiener University in Lima, Peru. His research interests are psychometrics and clinical and health psychology.

Ana Cecilia Quispe Álvarez, Master and professor at Universidad Católica San Pablo, Arequipa, Universidad Nacional San Agustín, Arequipa.

Adolfo Alexander Gamero Díaz, Master in Educational Sciences with a mention in Higher Education

Rosario Irma Butrón Ortiz, Specialist with a mention in initial education.

Mariel Delgado-Campusano, Psychologist from the Peruvian University of Applied Sciences. Researcher at the Peruvian University of Applied Sciences

Mario Reyes-Bossio, Ordinary Professor and Full-Time Research Professor at the Peruvian University of Applied Sciences (UPC, Peru). Teacher in undergraduate and postgraduate nationally and internationally. Sports Psychology and Physical Activity, Research Methodology, Psychometrics, Education and Organization.

**Corresponding Author’s Contact Address**

Universidad Científica del Sur, Lima, Perú.

Email: tcaycho@cientifica.edu.pe