Psychometric analysis and normative data of UWES in Peruvian adolescents

Análisis psicométrico y datos normativos de la UWES en adolescentes peruanos

Análise psicométrica e dados normativos da UWES em adolescentes peruanos

Sergio Dominguez-Lara¹, ORCID 0000-0002-2083-4278
Benigno Peceros-Pinto², ORCID 0000-0002-4865-3874
Sharon Centeno-Leyva³, ORCID 0000-0001-6827-6749
Sabina N. Valente⁴, ORCID 0000-0003-2314-3744
Abílio A. Lourenço⁵, ORCID 0000-0001-6920-0412
Alberto Quistgaard-Alvarez⁶, ORCID 0000-0002-1731-6787
Mercedes Patricia Morales-Velasquez⁷, ORCID 0000-0002-3135-4464

¹ Universidad de San Martín de Porres, Peru
² Universidad de San Martín de Porres, Peru
³ Innova Scientific, Peru
⁴ University of Évora, Portugal - Polytechnic Institute of Portalegre, Portugal
⁵ Universidade do Minho, Portugal
⁶ Universidad Privada San Juan Bautista, Peru
⁷ Universidad Privada San Juan Bautista, Peru

Abstract

Academic engagement is a relevant characteristic to predict successful academic trajectories and school performance; however, there is a lack of validated instruments in Peru for its evaluation. Therefore, the objective of this study was to analyze the psychometric properties of the Utrecht Work Engagement Scale – 9S (UWES-9S) in Peruvian adolescents: internal structure, measurement invariance, association with academic self-efficacy, reliability, and norms. A total of 868 school adolescents (51.728% women; M_age = 14.263; SD_age = 1.430) from six schools in the Constitutional Province of Callao (central coast of Peru) were evaluated. In addition to the UWES-9S, the Specific Perceived Self-Efficacy Scale for Academic Situations was used as a measure of self-efficacy. The results show that the UWES-9S is unidimensional and invariant between men and women, although three items were eliminated, resulting in a new version: UWES-6S. Likewise, the association with academic self-efficacy was moderate (r > .50) and norms were obtained for men and women separately. In addition, the reliability indices were satisfactory for both the scores (alpha coefficient and average inter-item correlation) and the construct (omega coefficient > .80). It is concluded that the UWES-6S has adequate psychometric properties for its application in Peruvian adolescents.

Keywords: academic engagement; adolescents; validity; reliability; test norms

Resumen

El engagement académico es una característica relevante para predecir las trayectorias académicas exitosas y el rendimiento escolar; sin embargo, existe una carencia de instrumentos validados en Perú para su evaluación. Por ello, el objetivo de esta investigación fue analizar las propiedades psicométricas de la Utrech Work Engagement
Educational changes due to COVID-19 pandemic affected millions of students from all education levels worldwide (Moreta-Herrera et al., 2022) and represented an unprecedented challenge to the education system (Tarabini, 2020) due to the abrupt change in its usual planning of the teaching processes’ input, development and completion to a remote system. This meant reduced class time, lower learning content, and unequal educational use due to the economic crisis’ impact on students’ homes (Save
the Children, 2020). Likewise, there were limitations on the families’ ability to accompany distance learning due parents’ work and the caregivers’ unfamiliarity with the proposed new system, as well as the effectiveness of a virtual environment that favors student learning (Failache et al., 2020). Inequalities became more evident according to socioeconomic level, educational institution’s management, place of residence, region, and/or locality (Cabrera et al., 2020) reducing the possibilities of access to technological resources and connectivity (Valente, 2020). This could increase school dropout rates (Lopera-Zuluaga, 2020).

At the beginning of the pandemic at least one third of students worldwide did not have access to distance education, leaving school (United Nations Children's Fund [UNICEF], 2020), and the costs are now estimated to be high regarding loss of learning, health, and increased school dropout. Thus, 258 million children were left out of school during the closure (United Nations Educational, Scientific and Cultural Organization [UNESCO], 2020), while two million girls and 5.7 million primary and secondary school children are at risk of leaving school (UNESCO, 2022).

Before the pandemic, school failure figures (not attending school, leaving school, or failing to achieve educational goals) in Peru were not encouraging (28.5% on the coast region; 32% in the mountainous region; 41% in the amazon region; Latorre, 2018), and recently school dropout situation in Peru (Ministerio de Educación del Perú [MINEDU], 2020) at the primary level increased from 1.3% to 3.5% (128,000 students), in secondary from 3.5% to 4% (102,000 students), and 337,870 students moved from private to public schools. On the other hand, by the end of 2021, around 83,000 students left school, and near 301,000 students did not access distance education (Defensoría del Pueblo, 2021). Therefore, school dropout is not a sudden act, but a final stage of a cumulative process of loss of interest and commitment to studies, as well as external triggers (Lara et al., 2018).

In this way, several researchers agree that academic engagement (AE) is a key concept to understand and prevent school dropout, as well as a significant predictor of the variables involved in the success of educational trajectories (Saracostti et al., 2019), including academic performance and attendance (Miranda-Zapata et al., 2018).

The AE is the degree in which students are involved in their learning (Chase et al., 2014; Usán-Supervía et al., 2018) and motivated to actively participate in their academic activities and achievements (Tomás et al., 2016). Furthermore, it is considered an influential variable in the advancement of student competences and an important predictor of academic results due to its association with other personal variables that also affect the school environment (Siu et al., 2014; Sulea et al., 2015).

In this way, some authors conceptualize it in a multidimensional way, either as a function of behavioral components (e.g. participation in class, attendance, performing school tasks), and components associated with cognitive affective-emotional commitment like sense of belonging, will, effort, enthusiasm, use of strategies, determination and self-regulation (Fredricks et al., 2016), as well as around three dimensions such as dedication, which refers to enthusiasm and joy against academic challenges, vigor associated with high levels of energy, and absorption or degree of concentration or immersion in activities (Schaufeli & Salanova, 2007).

The AE is associated with several variables, among which academic self-efficacy (ASE) stands out, which is defined as the students’ beliefs about their learning abilities (Khan, 2013) and considered as AE’ engine (Salanova et al., 2005) as well as an important variable in adolescence (Nunes & Faro, 2021). ASE is a key factor in the development of AE because success in previous activities will increase positive beliefs about the possibility of having a good academic performance and, consequently, the commitment to the activity (Medrano et al., 2015). In the field of basic education there is evidence of
direct association between the AE and the ASE (Olivier et al., 2019; Sağkal & Sönmez 2021; Usán-Supervía et al., 2018) Therefore, students who feel able to solve problems and understand concepts have more opportunities to get involved in didactic activities, whether in person or not, since they will feel they have necessary tools for it.

Regarding measuring AE in Peruvian schoolchildren, although there is a scale developed for Ibero-America and with acceptable psychometric properties (Lara et al., 2021), the in-person approach to items would not be appropriate for use in pandemic situations where education was not in-person. For example, for affective engagement, which is focused on how the student feels at school, some items focus on specific situations at school (e.g., "What we do at school is very important to me" or "I can be myself in this school") as well as behavioral engagement, which deals with behavior and discipline in the classroom (e.g., "I leave without asking permission from the classroom" or "I behave well in classes"), while cognitive engagement would be appropriate because it focuses on student motivation (e.g., "When I’m doing some activity, I worry about understanding as much as possible" or "When I study, I write down new words, doubts or important ideas"). In that sense, adapting the scale to a distance system would involve completely modifying it and even eliminating inappropriate items.

Therefore, the Utrech Work Engagement Scale - Student (UWES-9S; Schaufeli et al., 2006) was selected, which focuses on the pleasure of study activity, regardless of physical space or modality. In addition, the basic theory is articulated to a wide amount of research and has empirical support in different contexts, which allows us to know those difficulties around student performance, as well as school dropout (Salanova et al., 2005).

Finally, it is short (9 items), which allows optimal time management considering schools’ schedule. However, despite psychometric evidence in university students, psychometric studies in basic education students are scarce. Specifically, it has two studies in Spain (García-Ros et al., 2018; Serrano et al., 2019) and one in Finland (Salmela-Aro & Upadaya, 2012) which indicates the presence of a single factor in addition to being invariant according to sex.

The current study

There is little attention to the analysis of UWES-9S’s psychometric properties in schoolchildren, especially in Spanish-speaking countries in Latin America, although it is widely studied in university students in those countries. For this reason, the purpose of the present study was to analyze the psychometric properties of UWES-9S in Peruvian adolescents and thus, to fill this knowledge gap, since although it is mentioned that there is evidence that the UWES-S is unidimensional and invariant between men and women (García-Ros et al., 2018; Serrano et al., 2019), as far as the authors know, there are no published studies on its psychometric benefits or regulatory data (scales) in Peruvian adolescents.

This study is important since a tool that evaluates AE with evidence of validity and reliability will allow us to accurately assess the effectiveness of interventions in this variable (Leyton et al., 2021), considering its relevance in the student’s academic life.

Moreover, an invariance study is important, since women have better reading performance than men, but 62% more likely to have poor performance in science and mathematics (Muelle, 2018), so it is likely to influence the sex gap against women on access and completion of studies (Fuentes & Sánchez, 2010), which could account for an unfavorable scenario for women and, consequently, a biased assessment.

On the other hand, it is also useful that the scale studied be short, as this will allow the evaluation time to be optimized by not tiring the respondent and achieving more
reliable answers. It could be even easily inserted into evaluation protocols of broader studies in order to include the AE in models explaining academic performance (e.g., Miranda-Zapata et al., 2018).

Method

Design
This is an instrumental study (Ato et al., 2013), which studies the psychometric properties of UWES-9S in Peruvian adolescents.

Participants
Eight hundred and sixty-eight school adolescents (51.7% women) from six jointly managed private schools (State and Church) of Callao Constitutional Province (central coast of Peru), aged between 12 and 17 years ($M_{age} = 14.26; SD_{age} = 1.43$), mostly of medium-low socioeconomic level.

Instruments
Utrech Work Engagement Scale - Student (UWES-9S; Schaufeli et al., 2006). UWES-9S originally evaluates the three dimensions of AE (dedication, vigor and absorption) by 9 items (3 items per factor) in a Likert scale ranging from 0 (never) to 6 (always). The version’s content of items validated in Peru (Domínguez-Lara et al., 2021; Domínguez-Lara et al., 2020) was adapted to a basic education context.

Scale of Specific Perceived Self-Efficacy of Academic Situations (EAPESA, in Spanish; Palenzuela, 1983). The version adapted to Peruvian adolescents (Navarro-Loli & Domínguez-Lara, 2019) evaluates in a one-dimensional way the 7-item ASE with four options ranging from 1 (never) to 4 (always). Together with the data of this study, reliability coefficients of acceptable magnitude were found ($\alpha = .89$).

Procedure for collecting information
After approving the research project, several educational institutions were contacted. Initially, parents were informed of the research’s objectives and permission to evaluate their children was sought through informed consent. Survey administration to adolescents was carried out during class hours, and only those who agreed to participate after explaining the study’s objectives were evaluated in a more understandable language. Both informed consent for parents like self-reporting administration was done via a Google Forms link.

Preliminary to the instruments’ administration, some adjustments were made to the items in order to make them compatible with the age group evaluated following the recommendations of preliminary studies (e.g., García-Ros et al., 2018), which can be seen below (Table 1), and that version was evaluated by a small group of students ($n = 20; 50$% women) to verify if the content was comprehensible, and none of the participants expressed difficulties in understanding the items or in the UWES-9S response format.
Table 1
Adjustment of UWES-9S items to the school context

<table>
<thead>
<tr>
<th>University version</th>
<th>School version</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. My homework as a student makes me feel full of energy</td>
<td>1. My homework as a student makes me feel full of energy.</td>
</tr>
<tr>
<td>2. I feel strong and vigorous when I’m studying or going to classes</td>
<td>2. I feel strong and vigorous when I’m studying or attending to classes</td>
</tr>
<tr>
<td>3. I’m excited about my career</td>
<td>3. I’m excited about my studies</td>
</tr>
<tr>
<td>4. My studies inspire me new things</td>
<td>4. My studies inspire me new things</td>
</tr>
<tr>
<td>5. When I wake up in the morning, I feel like going to class to study</td>
<td>5. When I wake up in the morning, I feel like going to class to study</td>
</tr>
<tr>
<td>6. I am happy when I am doing homework related to my studies</td>
<td>6. I am happy when I am doing homework</td>
</tr>
<tr>
<td>7. I’m proud to follow this career</td>
<td>7. I’m proud to study</td>
</tr>
<tr>
<td>8. I am immersed in my studies</td>
<td>8. I am very focus on my studies</td>
</tr>
<tr>
<td>9. I &quot;get carried away&quot; when I do my homework as a student</td>
<td>9. I &quot;get carried away&quot; when I do my homework</td>
</tr>
</tbody>
</table>

Ethical considerations

This report is part of a larger research project approved by the Institutional Ethics Committee of Universidad Privada San Juan Bautista (Registration No. 063-2021-CIEI-UPSJB), and was conducted under the declaration of Helsinki and the ethics code of the Peruvian Association of Psychologists (2017).

Data analysis

Preliminary to the main analyses, univariate normality was evaluated through asymmetry magnitude ($g_1 < 3$; Kline, 2016) and kurtosis ($g_2 < 10$; Kline, 2016) of each item, and multivariate normality with Mardia’s G2 coefficient ($< 70$) was complementary. Cut-off points established to achieve an approximation to normality are conservative (Kline, 2016) because they were based on populations from studies in mental health and substance abuse, where the deviation from statistical normality was severe (Curran et al., 1996). Thus, in the context of item analysis, an asymmetric distribution would reaffirm the use of polychoric matrices (Ferrando & Lorenzo-Seva, 2019).

In the first stage of the validity evidence’s analysis in relation to internal structure, two UWES-9S measurement models were evaluated in men and women separately, the original of three factors (Schaufeli et al., 2006) and a one-dimensional model based on other studies published in Peru (Dominguez-Lara et al., 2021; Dominguez-Lara et al., 2020) and with Spanish adolescents (García-Ros et al., 2018; Serrano et al., 2019). The estimation method by Weighted Least Square Mean and Variance Adjusted (WLSMV) was implemented because it is an appropriate method for ordinal items such as those of the UWES-S (Ledesma et al., 2021; Verdam et al., 2016), in addition to allowing a more accurate estimation of factorial loads compared to other methods based on maximum likelihood (Li, 2016a, 2016b); and based on the matrix of inter-item polychoric correlations. Regarding the models’ psychometric strength, three criteria were considered: the adjustment indices by means of CFI’s magnitude ($>.90$), the upper limit of RMSEA’s confidence interval ($<.10$), and the WRMR ($< 1$), as well as the factorial loads’ magnitude ($>.50$) and the correlation between factors where values above .90 suggest multicollinearity (Brown, 2015).

The second stage consisted of an analysis of measurement invariance according to sex under a multigroup factor analysis approach with the best model between the two
proposed (three factors and unifactorial). Specifically, the invariance of the general structure (configural), of the factorial loads (metric), of the intercepts or thresholds (strong), and of the residuals (strict) was analyzed. An acceptable degree of invariance was determined depending on the CFI and RMSEA’s variation (ΔCFI > -.01; ΔRMSEA ≤ .015) from one stage to another (Chen, 2007).

Finally, the combined total sample (men and women) was divided in half to evaluate the one-dimensional model and refine the scale. The one-dimensional model was analyzed with the first group (M1), complemented by the progressive implementation of the correlations among residuals associated with the modification index with greater magnitude in order to detect those problematic items to improve the scale’s metric quality (Navarro-Loli & Dominguez-Lara, 2019; Pérez-Fuentes et al., 2020) and achieve a version that does not depend on the correlation among residuals to obtain adjustment indices or artificially elevated construct reliability magnitudes (Dominguez-Lara, 2019). With the second group (M2) the resulting version was analyzed and consolidated into M1. The assessment criteria were similar to the first stage.

The factor analyses were run with 7.0 version Mplus software (Muthén & Muthén, 1998-2015).

Regarding the validity evidence due to its relationship with other variables, the UWES score was correlated with a measure of academic self-efficacy. The correlation coefficient’s magnitude was used as an assessment criterion, where a value between .20 and .50 is considered low, between .50 and .80, moderate, and greater than .80, high.

After that, the reliability of the scores (alpha coefficient) and the construct (omega coefficient) were analyzed. Regarding the scores’ reliability, prior to the estimation of the alpha coefficient, the tau-equivalence (statistical equality of factor loads) within the scale was evaluated, and its contrast with the congeneric model (one-dimensional model without restrictions) that was carried out using the above-mentioned adjustment indices criteria. So, with respect to its magnitude, it was expected at least .70. Likewise, the average inter-item correlation ($r_{ij} > .40$) was considered as scores’ additional reliability criterion because, unlike the alpha coefficient, its magnitude is not dependent on the number of items (Clark & Watson, 1995; Raykov, 2012), since always a shorter version of a scale will present a lower alpha coefficient even if the items of both versions have the same strength of association (Domínguez-Lara et al., 2021). Subsequently, the construct reliability was evaluated with the omega coefficient ($> .80$). Similarly, the discrepancy between these was assessed descriptively ($\omega - \alpha$) and considered significant if it is greater than |.06| (Gignac et al., 2007). Then, with the definitive version in M1 and M2, both coefficients were corrected by the presence of correlated residuals.

Finally, after analyzing the differences between men and women, normative data (norms) were elaborated separately considering different levels based on percentiles (Pc): very low level (< Pc 10), low (< Pc 25), low average (between Pc 25 and Pc 50), high average (between Pc 25 and Pc 50), high (> Pc 75), and very high (> Pc 90). (Domínguez-Lara et al., 2022).

**Results**

**Descriptive analysis**

All items have acceptable magnitudes of asymmetry and kurtosis (Table 2), indicating a reasonable approximation to univariate normality. On the other hand, in terms of multivariate normality, there is favorable evidence for both groups ($G_{2men} = 17,156; G_{2women} = 23,387$).
Table 2
*UWES-9S: Items descriptive analysis*

<table>
<thead>
<tr>
<th>Item</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>M</em></td>
<td><em>SD</em></td>
<td><em>g1</em></td>
<td><em>g2</em></td>
</tr>
<tr>
<td>Item 1</td>
<td>2.759</td>
<td>1.600</td>
<td>0.233</td>
<td>-0.532</td>
</tr>
<tr>
<td>Item 2</td>
<td>2.905</td>
<td>1.637</td>
<td>0.203</td>
<td>-0.794</td>
</tr>
<tr>
<td>Item 3</td>
<td>3.382</td>
<td>1.519</td>
<td>0.027</td>
<td>-0.587</td>
</tr>
<tr>
<td>Item 4</td>
<td>3.704</td>
<td>1.587</td>
<td>-0.130</td>
<td>-0.820</td>
</tr>
<tr>
<td>Item 5</td>
<td>2.747</td>
<td>1.675</td>
<td>0.190</td>
<td>-0.760</td>
</tr>
<tr>
<td>Item 6</td>
<td>2.752</td>
<td>1.562</td>
<td>0.304</td>
<td>-0.472</td>
</tr>
<tr>
<td>Item 7</td>
<td>4.332</td>
<td>1.478</td>
<td>-0.424</td>
<td>-0.749</td>
</tr>
<tr>
<td>Item 8</td>
<td>3.842</td>
<td>1.436</td>
<td>-0.121</td>
<td>-0.649</td>
</tr>
<tr>
<td>Item 9</td>
<td>3.146</td>
<td>1.506</td>
<td>0.143</td>
<td>-0.617</td>
</tr>
</tbody>
</table>

Note. *M*: Mean; *SD*: Standard Deviation; *g1*: asymmetry; *g2*: kurtosis

Evidence of validity of the internal structure and invariance of UWES-9S in men and women

Regarding the internal structure, the two models (one and three factors) have similar adjustment indices (Table 3), but the average interfactorial correlation (φ) in both groups suggests a one-dimensional structure (φ<sub>women</sub> = .986; φ<sub>men</sub> = .980). In this way, the measurement invariance was analyzed, and the one-dimensional model was practically equivalent between men and women when considering the variability of the adjustment indices from one condition to another (Table 3), and even at the descriptive level with the magnitude of the factor loadings (Table 4).

Table 3
*Models and measurement invariance of the UWES-9S in schoolchildren*

<table>
<thead>
<tr>
<th></th>
<th>CFI</th>
<th>ΔCFI</th>
<th>RMSEA</th>
<th>CI90%</th>
<th>ΔRMSEA</th>
<th>WRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>.977</td>
<td>.129</td>
<td>.112 - .146</td>
<td>.913</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>.982</td>
<td>.135</td>
<td>.119 - .152</td>
<td>.932</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unidimensional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>.976</td>
<td>.125</td>
<td>.109 - .141</td>
<td>.957</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>.974</td>
<td>.150</td>
<td>.135 - .166</td>
<td>1.115</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measurement invariance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Configural</td>
<td>.975</td>
<td>.138</td>
<td>.127 - .149</td>
<td>1.470</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metric</td>
<td>.994</td>
<td>0.019</td>
<td>.065</td>
<td>.054 - .077</td>
<td>-0.073</td>
<td>1.701</td>
</tr>
<tr>
<td>Strong</td>
<td>.986</td>
<td>-0.008</td>
<td>.073</td>
<td>.065 - .082</td>
<td>0.008</td>
<td>1.659</td>
</tr>
<tr>
<td>Strict</td>
<td>.988</td>
<td>0.002</td>
<td>.067</td>
<td>.059 - .075</td>
<td>-0.006</td>
<td>1.714</td>
</tr>
</tbody>
</table>
Table 4
Factor loadings of unidimensional models of the UWES-9S in schoolchildren

<table>
<thead>
<tr>
<th>Items</th>
<th>Men</th>
<th>Women</th>
<th>S1</th>
<th>S1 (sin 9)</th>
<th>S1 (without 4, 6 y 9)</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>.802</td>
<td>.804</td>
<td>.825</td>
<td>.827</td>
<td>.815</td>
<td>.776</td>
</tr>
<tr>
<td>Item 2</td>
<td>.836</td>
<td>.853</td>
<td>.87</td>
<td>.87</td>
<td>.881</td>
<td>.829</td>
</tr>
<tr>
<td>Item 3</td>
<td>.887</td>
<td>.893</td>
<td>.89</td>
<td>.893</td>
<td>.884</td>
<td>.884</td>
</tr>
<tr>
<td>Item 4</td>
<td>.833</td>
<td>.868</td>
<td>.85</td>
<td>.849</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Item 5</td>
<td>.756</td>
<td>.819</td>
<td>.817</td>
<td>.819</td>
<td>.804</td>
<td>.758</td>
</tr>
<tr>
<td>Item 6</td>
<td>.829</td>
<td>.875</td>
<td>.866</td>
<td>.864</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Item 7</td>
<td>.733</td>
<td>.790</td>
<td>.773</td>
<td>.768</td>
<td>.768</td>
<td>.754</td>
</tr>
<tr>
<td>Item 8</td>
<td>.705</td>
<td>.722</td>
<td>.733</td>
<td>.734</td>
<td>.747</td>
<td>.698</td>
</tr>
<tr>
<td>Item 9</td>
<td>.422</td>
<td>.446</td>
<td>.417</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

\( \omega \)  
\( \alpha \)  
\( \omega - \alpha \)  

\( r_{ii} \)  
\( \text{Range}_{r_{ii}} \)  

Note. S1: Sample 1; S2: Sample 2; \( \omega \): omega coefficient; \( \alpha \): alpha coefficient; \( r_{ii} \): average inter-item correlation.

After the invariance analysis, the total sample was divided in half, and in the first subsample item 9 was eliminated according to the magnitude of its factorial load and the discrepancies with the other loads, both in men and women.

In this way, in the first subsample and with an 8-item version, the correlation between the residuals of items 8 and 7 was progressively modeled (Modification Index [MI] = 88.981; Estimated Parameter Change [EPC] = .407; \( \varphi = .371 \)), 1 and 6 (MI = 29.984; EPC = .305; \( \varphi = .267 \)), 5 and 6 (MI = 19.934; EPC = .250; \( \varphi = .227 \)), 2 and 4 (MI = 17.808; EPC = -.298; \( \varphi = -.336 \)), 4 and 6 (MI = 15.664; EPC = -.290; \( \varphi = -.320 \)), and between 1 and 4 (MI = 19.860; EPC = -.324; \( \varphi = -.368 \)), and items 4 and 6 were eliminated because they coincide in several of the correlations shown (Table 5). After that, the initial adjustment indices were similar to the 8-item version, and only the correlation implementation among the residuals of items 7 and 8 was necessary (MI = 96.158; EPC = .468; \( \varphi = .390 \)), so the resulting 6-item version (UWES-6S) had fewer poor specifications.

With the second subsample, the UWES-6S’ initial exploration showed more favorable indicators (Table 5), and subsequently the adjustment indices improved with the implementation of the correlation among the residuals of items 7 and 8 (MI = 35,074; EPC = .271; \( \varphi = .247 \)), as with the first sample.
Table 5  
*Measurement models in sample 1 and 2*

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model Description</th>
<th>CFI</th>
<th>RMSEA</th>
<th>CI90%</th>
<th>WRMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>Unidimensional</td>
<td>.976</td>
<td>.142</td>
<td>.127 - .158</td>
<td>1.099</td>
</tr>
<tr>
<td></td>
<td>Unidimensional (without item 9)</td>
<td>.981</td>
<td>.151</td>
<td>.134 - .170</td>
<td>1.115</td>
</tr>
<tr>
<td></td>
<td>Unidimensional (residuals specified)</td>
<td>.998</td>
<td>.065</td>
<td>.042 - .089</td>
<td>.441</td>
</tr>
<tr>
<td></td>
<td>Unidimensional (without 4, 6 and 9)*</td>
<td>.981</td>
<td>.171</td>
<td>.145 - .198</td>
<td>.999</td>
</tr>
<tr>
<td></td>
<td>Unidimensional (without 4, 6 and 9; correlation between residuals of items 7 and 8)</td>
<td>.996</td>
<td>.082</td>
<td>.053 - .113</td>
<td>.460</td>
</tr>
<tr>
<td></td>
<td>Unidimensional (without 4, 6 and 9/tau-equivalence)</td>
<td>.976</td>
<td>.155</td>
<td>.134 - .177</td>
<td>1.543</td>
</tr>
<tr>
<td>Sample 2</td>
<td>Unidimensional (without 4, 6 and 9)*</td>
<td>.985</td>
<td>.128</td>
<td>.102 - .156</td>
<td>.714</td>
</tr>
<tr>
<td></td>
<td>Unidimensional (without 4, 6 and 9; correlation between residuals of items 7 and 8)</td>
<td>.992</td>
<td>.100</td>
<td>.072 - .130</td>
<td>.513</td>
</tr>
<tr>
<td></td>
<td>Unidimensional (without 4, 6 and 9/tau-equivalence)</td>
<td>.974</td>
<td>.137</td>
<td>.116 - .159</td>
<td>1.398</td>
</tr>
</tbody>
</table>

*Compared with tau-equivalence model

**Validity evidence by its relation with other variables**

Association between UWES-6S and measure of academic self-efficacy was moderate ($r = .620$).

**Reliability**

Initially, tau-equivalence was analyzed before the estimation of the alpha coefficient, and it received favorable evidence in both the first subsample ($\Delta$CFI = -.005; $\Delta$RMSEA = -.016) and the second ($\Delta$CFI = -.011; $\Delta$RMSEA = .009).

Regarding the scores’ reliability, both the alpha coefficient and the $r_{ij}$ were acceptable in all cases, according to sex and between samples, and a similar picture was appreciated with the construct reliability (omega). Similarly, the difference between alpha and omega coefficients is negligible (Table 4). Finally, the correction of alpha and omega coefficients in the presence of correlated errors reflects a greater impact on the first one (Table 4).

**Normative data**

A small magnitude difference ($d = 0.305$) was found between men ($M = 19.967; SD = 7,512$) and women ($M = 17,648; SD = 7,681$), which led to the elaboration of separate scales (Table 6).
Table 6
Normative data according to the sex

<table>
<thead>
<tr>
<th>Level</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very low</td>
<td>Lower than 11</td>
<td>Lower than 8</td>
</tr>
<tr>
<td>Low</td>
<td>From 11 to 13</td>
<td>From 8 to 11</td>
</tr>
<tr>
<td>Average-Low</td>
<td>From 14 to 18</td>
<td>From 12 to 16</td>
</tr>
<tr>
<td>Average –High</td>
<td>From 19 to 25</td>
<td>From 17 to 21</td>
</tr>
<tr>
<td>High</td>
<td>From 26 to 30</td>
<td>From 22 to 28</td>
</tr>
<tr>
<td>Very high</td>
<td>31 or higher</td>
<td>29 or higher</td>
</tr>
</tbody>
</table>

Discussion

The objective was to determine if UWES-9S has psychometric properties suitable for use in Peruvian schoolchildren, as well as to elaborate regulatory data for a practical use of the scale in applied environments.

Regarding factor analysis, the one-dimensional structure presents better psychometric credentials than the multidimensional one due to the overlap of the three factors (correlations close to the unit), which was also observed by studies carried out in Spain (García-Ros et al., 2018; Serrano et al., 2019), in addition to being an invariant measure between men and women. In a complementary way, the correlation between residuals was analyzed, which helped to improve the metric quality of the instrument because it allowed to eliminate three items that shared sources of variance beyond the evaluated construct, which could interfere with the scores’ interpretation (Brown, 2015). Implementing correlations among residuals was carried out in a previous study (Serrano et al., 2019), but with the explicit objective of improving the adjustment of the model and without making corrections to the reliability coefficients, which could affect the interpretations’ validity (Dominguez-Lara, 2019), while in another study the items were grouped into plots (García-Ros et al., 2018), which accounts for the one-dimensional character and the need to further reduce the scale to improve their metric indicators. On the other hand, its association with a measure of academic self-efficacy was corroborated, confirming what was established by the specialized literature (Salanova et al., 2005).

As for reliability, tau-equivalence was analyzed prior to the estimation of the coefficients. The indicators are favorable at both the score and construct level (≈ .90), allowing for more accurate use in both individual and group applications, including the use of the UWES-6S in explanatory models involving the study of latent variables. However, although after correction for the presence of correlated errors the reliability coefficients decrease, it is necessary to know that the magnitude of the correlations between residuals (which directly affect the corrected magnitude) varies according to the sample evaluated, so they must be considered each time the scale is used.

As for the regulatory data, it was found that men score higher than women in AE, which reaffirms the disadvantages accentuated during the pandemic.

Regarding the practical implications, the results support the solidity AE’s short scale in Peruvian adolescents, which will minimize the application time during its use by reducing the fatigue of the respondent and generating valid responses since, due to the characteristics of the sample, there are difficulties in understanding instructions and in establishing judgments about one's own behavior. It will also help design appropriate educational intervention programs that focus on adolescent AE development. In this way, possibilities are opened in the field of basic research since it can be used in broader projects simultaneously with other instruments and thus have information on various
aspects of behavior, as well as in cross-cultural studies that compare AE in Spanish-speaking adolescents from different contexts.

As for work’s strengths, there was a large sample of several educational institutions, which diversifies the responses and minimizes the bias associated with a single educational institution. This allowed to have a replication sample and thus have more consistent evidence on the new internal structure (UWES-6S). In the same way, the multiple sources of information (factor analysis, relations with other variables, invariance, and reliability) allow us to conclude with greater evidence about its one-dimensional structure. However, there are some limitations, since, despite the number of participants, the sample was non-probabilistic, that is, it does not adequately represent all students in secondary education. Likewise, the influence of social desirability in the evaluated group is likely (Fariña-Rivera et al., 2021), and the use of various self-report instruments within the same evaluation protocol could attenuate the psychometric strength of the scales due to external sources of error, although despite these limitations, the UWES-6S provides valuable scores to measure AE in schooled adolescents.

Therefore, it is concluded that UWES-6S has adequate psychometric properties in Peruvian schoolchildren. Specifically, it presents a solid factorial structure with factorial loads of magnitudes ranging from moderate to high; it has, in turn, evidence of invariance between men and women; and is theoretically consistently associated with academic self-efficacy. As for reliability, its magnitude favors its use in basic and applied research environments, which together with the use of regulatory data will allow obtaining relevant information to monitor intervention programs, whether individual or group.

It is recommended to replicate the procedures in a similar context after a while in order to assess the UWES-6S’s psychometric performance, because being data collected in a specific period (pandemic), the results may be focused on aspects of the predominant study modality (not face-to-face). It is also necessary to analyze the UWES-6S’s relevance in other educational settings, such as publicly managed schools since this will minimize the bias associated with the initial context of application (Merrell, 2000).

On the other hand, it would be interesting to examine the UWES-6S’s relation with other demographic variables such as origin region or educational level, as well as other psychological variables like academic exhaustion, psychological capital or academic performance. Additionally, it is advisable to deepen the AE’s determinants study according to sex, since the evidence indicates that women score below men, in contrast to other studies where differences are not appreciated (e.g., García-Ros et al., 2018), as well as their social determinants, given that there is a family gap that generates obstacles not only with respect to the lack of electronic devices, but also to a lower education of families, which brings with it greater difficulties in helping with homework at home; becoming decisive factors for educational outcomes and a possible increase in school dropout rates (Lopera-Zuluaga, 2020).

References


issues. Learning and Instruction, 43(1), 1-4. https://doi.org/10.1016/j.learninstruc.2016.02.002


**Authors’ participation:** a) Conception and design of the work; b) Data acquisition; c) Analysis and interpretation of data; d) Writing of the manuscript; e) Critical review of the manuscript.

S. D. L. has contributed in a, c, d, e; B. P. P. in b, d, e; S. C. L. in a, d, e; S. N. V. in d, e; A. A. L. in d, e; A. Q. A. in d, e; M. P. M. V. in d, e.

**Managing scientific editor:** Dra. Cecilia Cracco.