Prediction of disruptive behaviors from boredom and satisfaction with physical education

Predicción de las conductas disruptivas a partir del aburrimiento y la satisfacción con la educación física

***Raúl Baños*[1]**
***Michelle Barretos-Ruvalcaba*[1]**
***Antonio Baena-Extremera*[2]**
***Antonio Granero-Gallegos*[3]**

[1] Facultad de Deporte, Universidad Autónoma de Baja California, Mexico
[2] Facultad de Ciencias de la Educación, Universidad de Granada, Spain
[3] Facultad de Ciencias de la Educación, Universidad de Almería, Spain

Correspondence:
Raúl Baños
raulfb89@gmail.com

Short title:
Prediction of disruptive behaviors from boredom

How to cite this article:
Baños, R., Barretos-Ruvalcaba, M., Baena-Extremera, A., & Granero-Gallegos, A. (2024). Prediction of disruptive behaviors from boredom and satisfaction with physical education. *Cultura, Ciencia y Deporte, 19*(59), 89-101. https://doi.org/10.12800/ccd.v19i59.2173

Received: 28 December 2023 / **Accepted**: 07 February 2024

Abstract

The aim of the present study was to analyze the psychometric properties of the Questionnaire for Disruptive Behavior in Physical Education (CCD-EF) in the Mexican context. A non-experimental, cross-sectional, correlational-causal study was designed in which 378 girls (*M* = 13.99; *SD* = .30) and 375 boys (*M* = 14.02; *SD* = .33), all high school students participated. The psychometric properties of the scale were analyzed by means of different exploratory and confirmatory analyzes that demonstrate that this instrument with four correlated factors, and as higher order models, is valid, reliable, and invariant as a function of sex. A regression model with latent variables showed a positive and significant prediction of boredom with Physical Education on disruptive behaviors, finding that this prediction is higher in boys than in girls. The CCD-EF has proven to be a reliable and valid instrument to use with Mexican high school students.

Keywords: Physical education, disruptive behaviors, boredom, satisfaction, high school.

Resumen

El objetivo del presente estudio fue analizar las propiedades psicométricas del Cuestionario de Conducta Disruptiva en Educación Física (CCD-EF) en el contexto mexicano. Se diseñó un estudio no experimental, transversal, correlacional-causal, en el que participaron 378 niñas (*M* = 13.99; *DT* = .30) y 375 niños (*M* = 14.02; *DT* = .33), todos estudiantes de secundaria. Las propiedades psicométricas de la escala fueron analizadas mediante diferentes análisis exploratorios y confirmatorios que demuestran que este instrumento con cuatro factores correlacionados, y como modelos de orden superior, es válido, confiable e invariante en función del sexo. Un modelo de regresión con variables latentes mostró una predicción positiva y significativa del aburrimiento con la Educación Física sobre las conductas disruptivas, encontrando que esta predicción es mayor en niños que en niñas. El CCD-EF ha demostrado ser un instrumento confiable y válido para utilizar con estudiantes mexicanos de secundaria.

Palabras clave: Educación física, conductas disruptivas, aburrimiento, satisfacción, secundaria.

Introduction

School violence is one of the most worrying social problems in adolescents, and it leads to behaviors such as physical and verbal abuse among students, with disrespect for the teacher's authority, constant disorder in the classroom being among the most common (Manzano, 2021). In addition to this, classroom violence increases the probability of school failure, which, together with the dissatisfaction of adolescents, is even related to a higher suicide rate (Benbenisthy et al., 2018), which increases concern about this problem (Hinduja & Patchin, 2015). In fact, interpersonal violence is the fourth leading cause of adolescent and youth mortality worldwide, although its weight varies substantially from one region of the world to another, according to the World Health Organization [WHO] (WHO, 2021). In the low- and middle-income countries of the Region of the Americas, it is the cause of almost one third of all adolescent male deaths. According to data from the Global School Health Survey (GSHS), 42% of adolescent boys and 37% of girls were exposed to bullying (WHO, 2021), with low school engagement and school failure being a major risk factor (WHO, 2020). Different studies (Baños, Baena-Extremera & Granero-Gallegos, 2019; Fernández-Garcimartín et al., 2023; Hoyo-Guillot & Ruiz-Montero, 2023; Palacios-Gómez et al., 2023), have highlighted the importance of studying the disruptive behaviors of young people at school since they spend most of their time there.

These matters are of even greater concern in the Mexican context, as it continues to rank first internationally among Organization for Economic Cooperation and Development [OECD] countries in cases of bullying and school violence, a phenomenon that affects more than 18 million primary and secondary students in public and private schools in the country (Senado de la República de México, 2019). Furthermore, Mexico is one of the three OECD member countries where suicide rates rose between 1990 and 2015 for young people aged 15 to 19 years (Burns & Gottschalk, 2019).

Disruptive behaviors are defined as those that disrupt school activities, particularly behaviors that disturb teachers and/or students in class (Maddeh et al., 2015). For teachers, this set of inappropriate student behaviors hinders the normal development of classroom activities during the sessions, causing a great loss of time due to constant calls for attention that cause interruption of school planning and low academic performance (Sulbarán & León, 2014). Regarding the typification of disruption, school regulations and teachers› perceptions classify it under misconduct as: misbehavior, aggressiveness, repeated annoyance, indiscipline, rudeness, transgression of the truth, attention deficit and distancing (Sulbarán & León, 2014).

The subject of Physical Education (PE) has been labeled as a conflictive curricular area due to the frequent interactions that occur in it because of its eminently practical and competitive nature (Buscà et al., 2014). These conflicts in the PE classroom are usually more prone to occur in sessions where students have to demonstrate that they are more skilled than their peers (Baños, 2021), when students find PE classes monotonous and boring (Olweus & Breivik, 2014), and when they experience feelings of dissatisfaction with the subject (Baños, 2020). By contrast, when the adolescent experiences feelings of fun and satisfaction with PE the probability of negative behaviors in the classroom is reduced. Thus, boredom is a determining factor in the increase of inappropriate behaviors in PE (Baños et al., 2017), with the teacher being an important element in reducing it (Baños, Barretos-Ruvalcaba et al., 2019).

There are certain behaviors that are identified more frequently in PE, such as aggressive, low engagement or irresponsibility, fails to follow directions, distract or disturbs others, and poor self-management (Cothran & Kulinna, 2007; Kulina et al., 2006). Among the students' explanations for their inappropriate behaviors in the PE classroom the boredom they experience in their classes stands out as they find them to be monotonous, and dissatisfaction with the teacher is also significant. It is common to find these behaviors reiterative in the same subjects (Cothran & Kulinna, 2007). Along these lines, disruptive behaviors in the PE class can reduce learning and the opportunity for classroom activities (Supaporn et al., 2003). As some studies highlight, adolescents often show some problematic behaviors in classrooms, such as idleness, disrespect, irresponsibility, disobedience of rules, talking out of turn, and/or skipping class, which generates a negative impact on the learning environment (Baena-Extremera et al., 2015; Kulinna et al., 2006). The learning environment that the teacher is able to create in PE classes can influence inappropriate behaviors, motivation, and student learning (Baños et al., 2017).

In this vein, questionnaires have been designed and validated to measure disruptive behaviors in the PE classroom from the students' (Granero-Gallegos & Baena-Extremera, 2016; Krech et al., 2010) and the teachers' perspectives (Krech et al., 2010; Kulinna et al., 2006). Focusing on students, first Kulinna et al. (2003) and Cothran and Kulinna (2007) designed the 59 item Physical Education Classroom Instrument (PECI) to measure both the frequency of indiscipline and the severity of students' disruptive behavior. This scale began with six factors: aggressive, low engagement or irresponsibility, fails to follow directions, distract or disturbs others, poor self-management, and disruptive and illegal behaviors (Kulinna et al., 2003). However, none of these studies presented validity analyses of the instrument and also did not provide values related to goodness-of-fit indices. Later, the PECI evolved to a short version of 20 items from which the dimension of harmful and illegal behaviors was eliminated, and which did present acceptable validity and reliability values, Krech et al. (2010). Subsequently, the PECI was adapted and validated for PE in the Spanish context by

Granero-Gallegos and Baena-Extremera (2016) receiving the name Questionnaire for Disruptive Behavior in Physical Education (CCD-EF). This adaptation obtained acceptable validity and reliability values. However, in Mexico there are no adapted and validated questionnaires available that assess disruptive behaviors in PE classrooms and that could contribute to evaluating the origin of the very high rates of violence among Mexican adolescents. Likewise, it is worth mentioning that none of the previously mentioned versions of this instrument have analyzed it as a higher-order model (Lévy-Mangin & Varela, 2006) in which negative behaviors are evaluated from a global perspective, in addition to evaluating their variables in a correlated manner. Given that previous research presented a high correlation between the different factors (Granero-Gallegos & Baena-Extremera, 2016; Krech et al., 2010), it would be interesting to analyze the psychometric properties of the CCD-EF as a higher-order model in which it would function as a single latent variable.

It should be noted that of all the studies mentioned above, the only one that reported results of factorial invariance as a function of academic level was the study by Krech et al. (2010), who studied factorial invariance as a function of academic level (primary vs. secondary) and as a function of student and teacher perception. However, in the adaptation and validation of the instrument to the Spanish context (Granero-Gallegos & Baena-Extremera, 2016) the factorial invariance of the instrument was not reported, although recently obtained valid results in the factorial invariance analysis of the CCD-EF as a function of gender with another Spanish sample (Martínez-Molina et al., 2020). However, it is necessary to continue studying the validity of the instrument in different samples.

In view of the above, adapting and validating this instrument to the Mexican context may be of great interest, as it could contribute to the analysis of violence among school adolescents with such worrying figures as discussed in the previous paragraphs. In this way, PE teachers could establish general strategies to address disruptive environments and maintain discipline in the classroom. Therefore, the objectives of the present study are: (i) to analyze the psychometric properties of the CCD-EF scale in the Mexican context; (ii) to analyze a higher order model from the subscales included in the CCD-EF; (iii) to analyze the factorial invariance of the CCD-EF scale as a function of gender; and (iv) to analyze the predictive relationship of satisfaction with PE on disruptive behaviors. Based on all the literature reviewed, the following hypotheses are proposed based on the objectives of the present study: H1, the CCD-EF will obtain adequate validity and reliability in the Mexican context; H2, the higher order model will obtain adequate validity and reliability values; H3, the CCD-EF will be gender invariant; H4, satisfaction with PE will negatively and significantly predict disruptive behaviors in the classroom; H5, boredom with PE will positively and significantly predict disruptive classroom behaviors. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) initiative (Von Elm et al., 2008) was used to description of the research.

Method

Design

The present study has a non-experimental, cross-sectional, and correlational-causal design (Hernández et al., 2014). The sample design was probabilistic by centers, stratified, multistage by proportional allocation, and was made up of students in the third year of secondary education in the state of Nuevo León (Mexico). The secondary schools that participated in the study were randomly selected. To do this, all schools were listed and selected using a random number table. The number of students in the third year of secondary school in the State of Nuevo León in 2019 was 13396 girls and 13831 boys. These data were provided by the Secretary of Public Education of the State of Nuevo León.

Eligibility criteria

Any secondary school in the State of Nuevo León could be selected to participate in the study. The main condition is that only public secondary schools would participate in the randomization process. Once all centers had been enumerated, they were randomly selected using a numerical randomization table. The criteria for excluding the schools from the present study were that they were private schools or that they received a negative response from the director of the secondary education center to participate in the study. In addition, the subjects who did not present the informed consent signed by the father/mother/guardian to participate in the research were excluded.

Adaptation to the Mexican context

The CCD-EF items from Granero-Gallegos and Baena-Extremera (2016) were adapted to the Mexican context. Two native speakers adapted the 20 items to the Mexican context. To judge the goodness of the adaptation, the degree of coincidence with the original version was considered. Following Lynn (1986), four experts in secondary education analyzed the final version to ensure the adequate design of the items for the measurement of the construct it was intended to measure and that it maintained the original meaning. The experts evaluated the relevance and comprehension of each item on a scale from 1 (Strongly Disagree) to 5 (Strongly Agree). If the mean item scores were < 2.5, the items were revised. If an item was not rated by at least three of the four experts within the theoretical dimensions of the scale, it was further revised. The overall agreement of four experts on relevance and comprehension was measured by the Intraclass Correlation Coefficient (*ICC*); the values obtained were: *ICC* = .83 for relevance and *ICC* = .85 for comprehension. The Mexican version was administered to 45 high school

students between 12 and 14 years of age and they showed full comprehension of the items. The final Mexican version of the CCD-EF was thus arrived at.

Instruments

The Any secondary school Questionnaire for Disruptive Behavior in PE (CCD-EF). The CCD-EF validated to the Spanish context by Granero-Gallegos and Baena-Extremera (2016) was adapted to the Mexican context from the original Physical Education Classroom Instrument by Krech et al. (2010). This instrument presents 20 items that measure negative behaviors in the PE classroom. It is composed of five dimensions with four items each: a) aggressive, b) low engagement or irresponsibility, c) fails to follow directions, d) distract or disturbs others, and e) poor self-management. The scale was preceded by the statement, "Tell us your degree of agreement or disagreement regarding PE classes." Responses were collected on a 5-point Likert- type scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

Satisfaction with PE (SSI-EF). The version adapted to the Mexican context by Baños, Baena-Extremera, and Granero-Gallegos (2019) of the Sport Satisfaction Instrument (SSI) by Castillo et al. (2001) was used. This scale is composed of eight items and measures levels of satisfaction/fun with PE (five items) and boredom with PE (three items). The scale was preceded by the statement: "Tell us your degree of agreement or disagreement with regard to the PE classes" The responses were collected using a Likert type scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

Procedure

The Any This research was conducted in accordance with the Declaration of Helsinki 1961 (revised in Edinburgh in 2000). Approval was obtained by the Secretaría de Educación Pública of México and Universidad Autónoma of Baja California, México (identification number: 431/569/E). In order to carry out this research, a research proposal was submitted to the Secretaría de Educación Pública de México entitled: "Programme for International Student Assessment: relationship of school performance in secondary school students with psychological, family and physical activity variables", which was approved and subsidized by the aforementioned organization. Subsequently, permission was requested from the management of the secondary schools, and the parents/guardians of the students gave informed consent after the objectives and purpose of the study were stated to them. Once the aforementioned permissions had been obtained, data collection proceeded by informing the participants of the objective of the study, the anonymous and voluntary nature of their participation, the confidential treatment of their responses, that there were no right or wrong answers and requesting that they answer with the utmost sincerity. The questionnaires were completed in the classroom, with the principal investigator always present to resolve any doubts during the process, which lasted 15-20 minutes.

Risk of bias

Regarding bias control, it should be noted that there was randomization of the educational centers but not of the participants. There was blinding between the participants and the investigators responsible for the treatment and data analysis. Regarding selection bias, participation in the study was voluntary and communication with the students was in person in the classroom.

Sample size

Any secondary school Regarding the sample size, the requirement stipulated by Carretero-Dios and Pérez (2005) were met; namely, that there were 10 participants for each item to carry out a Confirmatory Factor Analysis (CFA). In addition, the Free Statistics Calculator v.4.0 (Soper, 2022) software was used, which calculated that a minimum of 717 subjects would be sufficient to detect effect sizes ($f2$) = 0.19 with a statistical power of .95 and a significance level of α = .05 in a structural equation model (SEM) with eight latent variables and 28 observed variables. A total of 755 students participated in the research.

Statistical analysis

For the achievement of the objective of this study, statistical analyses of homogeneity of each item and Cronbach's alpha (α) were carried out in the first instance

The evaluation of the factorial structure of the instrument analyzed was carried out through confirmatory factor analysis (CFA) Several fit indices were calculated for the evaluation of the different proposed models, combining absolute and relative fit indexes (Bentler, 2007; Markland, 2007). Among the absolute ones: the *p*-value, associated with the Chi-square statistic ($\chi 2$), the ratio between and degrees of freedom (gl; $\chi 2/gl$) and *GFI* (goodness-of-fit index). Among the relative indexes: *TLI* (Tucker-Lewis index) and *CFI* (comparative fit index). Also, *RMSEA* (root mean square error of approximation) and *SRMR* (standardized root mean square), were used as incremental indexes (Browne & Cudeck, 1993). In the case of $\chi 2/gl$ model fit values < 5.0 were considered acceptable, and ratios < 2.0 considered as indicators of excellent model fit (Tabachnick & Fidell, 2007). Values of *CFI* and *TL I*> .95, *RMSEA* < .06 and *SRMR* < .05, were considered as indices of excellent model goodness-of-fit (Hooper et al., 2008; Hu & Bentler, 1998). Values of *CFI* and *TL I* > .90 and *RMSEA* and *SRMR* < .08 were considered acceptable indices of model goodness-of-fit (Hooper et al., 2008; Marsh et al., 2004). The reliability of each dimension was assessed with α, and composite reliability (*CR*) and average variance were extracted (*AVE*) to measure convergent validity. Reliability values > .70 and *AVE* > .50 were considered acceptable as were values > .50 (Hair et al., 2014).

Attending to the high correlations that often occur between lower-order factors, researchers follow their initial CFA with a post-hoc test of a higher-order (e.g., second order) model (Myers et al., 2014). In a higher-order model, each item is specified as loading on its target first-order

subscale and each first-order factor is allowed to load on one or more higher-order factors (Rindskopf & Rose, 1988). Thus, alternative models were tested to find the best factor structure approach of the CCD-EF, including a higher-order model (H-CFA) as an alternative to the first-order model (CFA).

The invariance according to the sex variable (multigroup invariance) of the CCD-EF was tested following the methodological proposal advanced by Milfont and Fisher (2010), of testing four successively more constrained models. Following Chen (2007), these nested models were compared considering the change (*Δ*) in goodness-of-fit indexes (i.e., increase in *RMSEA* of at least .015 or decrease in *CFI* and *TLI* of at least .01, indicate lack of invariance).

Finally, to test the nomological validity of the scale, a regression model with latent variables was performed to test the predictive relationship of the SSI-EF (satisfaction/enjoyment and boredom with PE). For a better interpretation of the results, R2 was used as effect size (*ES*), having calculated the *CI* (95%) to confirm that no value was below the minimum required for interpretation (*R2* = .02) (Domínguez-Lara, 2017). The cut-off points established to score it are: .02, .13 and .26, for small, medium, and large, respectively (Cohen, 1992).

Results

Participants

The representative sample was calculated according to sex, for a finite population with a confidence level of 95% and a margin of error of + 5%, 378 being girls (*Mage* = 13.99; *SD* =. 30) and 375 boys (*Mage* = 14.02; *SD* = .33).

Table 1. Descriptive, internal consistency and homogeneity statistics (n = 375)

Dimensions:	*M*	*SD*	*CCIT-c*	*α without item*	*Q1*	*Q2*
Factor 1: Aggressive (**α** = .56)						
1. Amenazo a los demás compañeros/as de clase. [Threatens others]	1.47	0.95	0.35	0.48	2.23	4.39
2. Hablo correctamente a mis compañeros/as de clase. [Smart mouth toward students]	2.27	1.15	0.38	0.45	0.65	- 0.32
3. Hablo de los demás a sus espaldas. [Talking back]	1.81	1.10	0.32	0.49	1.36	1.05
4. Hablo de forma correcta al profesor/a. [Smart mouth toward teacher]	1.90	1.17	0.32	0.50	1.34	0.91
Factor 2: Low Engagement or Irresponsibility (**α** = .69).						
5. Me quejo habitualmente. [Whining]	2.57	1.33	0.42	0.61	0.41	- 0.96
6. Soy perezoso en clase. [Lazy]	2.47	1.37	0.42	0.62	0.46	- 1.01
7. Busco llamar la atención. [Attention seeking]	1.73	1.09	0.45	0.59	1.43	1.08
8. Me muevo lentamente a propósito. [Moves slowly on purpose]	1.89	1.22	0.50	0.56	1.23	0.37
Factor 3: Fails to Follow Directions (**α** = .71).						
9. Interrumpo las clases. [Interrupts]	1.72	1.12	0.53	0.70	1.55	1.43
10. Me siento inseguro en clase de EF. [Unsafe actions]	1.96	1.29	0.45	0.74	1.10	- 0.06
11. No presto atención en clase de EF. [Doesn't pay attention]	2.07	1.37	0.57	0.67	1.02	- 0.30
12. No sigo las instrucciones. [Not following directions]	1.86	1.17	0.64	0.64	1.23	0.41
Factor 4: Distract or Disturbs Others (**α** = .85).						
13. Tengo mal genio y me enojo. [Temper tantrums]	2.08	1.26	0.48	0.83	0.92	- 0.30
14. Abandono el grupo durante una actividad. [Leaving the group during an activity]	1.69	1.11	0.69	0.72	1.68	1.90
15. Miento en clase. [Lying]	1.71	1.11	0.68	0.73	1.58	1.54
16. Me salto las clases de EF. [Sneaks out of class]	1.51	1.00	0.67	0.74	2.10	3.62
Factor 5: Poor Self-Management (**α** = .70)						
17. Soy peleónero/a. [Quarrelsome]	1.82	1.16	0.58	0.58	1.31	0.70
18. Me burlo de otros/as compañeros/as de clase. [Makes fun of other students]	1.89	1.18	0.58	0.57	1.22	0.45
19. Argumento mis actos. [Arguing]	2.76	1.39	0.32	0.75	0.17	- 1.19
20. Acoso a algunos/as compañeros/as de clase. [Bullying]	1.49	1.08	0.50	0.63	2.29	4.12

Note: *M* = Mean; *SD* = Standard Deviation; *CCIT-c* = Corrected coefficient of item-total correlation; **α** = Cronbach's alpha; *Q1* = Skewness; *Q2* = Kurtosis.

Item descriptive analysis

The item statistics for each of the dimensions are presented in Table 1. In the aggressive dimension, it should be noted that the *SD* of item 1 was < 1 and the skewness and kurtosis values presented values < 2. In addition to this, the internal consistency of the dimension with the four items was inadequate (α = .56) and does not improve if any of the items are eliminated. These results should be considered for the evaluation of model fit in confirmatory analyses. The items of the dimensions of low engagement or irresponsibility, fails to follow directions and distract or disturbs others, obtained adequate values in *SD*, *CCIT-c*, skewness, kurtosis, and reliability. The irresponsibility and low commitment factor obtained values slightly below those accepted in reliability; however, according to Taylor et al. (2008), when a factor is composed of a small number of items (in this case by four items) an internal consistency index < .70 can be considered acceptable. Regarding the items of the poor self-management factor, the deletion of item 19 improved the internal consistency to .75 and item 20 obtained skewness and kurtosis values > *2*. These results should be considered for the evaluation of model fit in confirmatory analyses.

Confirmatory factor analysis and reliability

The representative Next, the original dimensionality theoretically proposed by Krech et al. (2010) was analyzed with CFA and, following authors such as Markland (2007), several models were formulated and analyzed, given that the data so recommended, and the most relevant results were reported. Considering the above in the analysis of scale items, it was appropriate to perform and compare several structural regression models to check the best fit. Several models were hypothesized (see Table 2). The first of the models included the 20 items and the five factors of the original scale (Krech et al., 2010) that presented some unacceptable fits (*TLI* = .87; *CFI* = .89), and two of the items of the aggressive dimension presented low regression weights < .33 (Hair et al., 2014) and high values (> 2.58) in their standardized residuals (Byrne, 2013). Considering the above, as well as the unacceptable internal consistency of the aggressive factor (α = .56) this factor was eliminated and a model with the other four factors was evaluated, it presented acceptable goodness-of-fit indices. However, taking into account the index modification values of the statistical program, the errors of items 11 and 12 of the fails to follow directions factor and items 14 and 16 of the distract or disturbs others factor were correlated and the model presented acceptable (*TLI* = .94; *CFI* = .95) and excellent (*SRMR* = .036; *RMSEA* = .058) goodness-of-fit indices. Thus, H1 is fulfilled. Despite the good fits of this model, it was considered that the high correlations (.94) between some of the factors (i.e., distract or disturbs others with poor self-management and fails to follow directions with low poor self-management) could limit the discriminant validity of the scale, as values < .85 are considered adequate (Henseler et al., 2015), although some authors consider values <. 90 to be (Teo et al., 2008), so a higher-order model (Lévy-Mangin & Varela, 2006) was evaluated, which also presented acceptable (*TLI* = .94; *CFI* = .95) and excellent (*SRMR* = .038; *RMSEA* = .060) goodness-of-fit indices. To examine the model comparison, the AIC (Akaike Information Criteria) and BIC (Bayesian Information Criterion) were also taken into account, in which, although they do not describe the model fit, lower values are considered to reflect a better fit (Table 3).

Table 2. Fit indices for each model

Models	χ^2	*df*	*p*	χ^2/*gl*	*TLI*	*CFI*	*SRMR*	*RMSEA (90%IC)*	*AIC*	*BIC*
Model 5 factors	765.80	160	.000	4.79	.87	.89	.049	.071 (.066;.076)	865.80	1097.00
Models 4 factors	456.10	98	.000	4.65	.91	.92	.042	.070 (.063;.076)	532.10	707.81
Models 4 factors*	339.49	96	.000	3.54	.94	.95	.036	.058 (.051;.065)	419.49	604.45
Higher order model*	361.22	.98	.000	3.69	.94	.95	.038	.060 (.053;.066)	437.22	612.94

Note: $\chi 2$ = chi-square; *df* = degrees of freedom; *TLI* = Tucker Lewis index; *CFI* = comparative fit index; *SRMR* = Standardized Root Mean-Square; *RMSEA* = root-mean squared approximation; *IC* = confidence interval; *AIC* = Akaike Information Criteria; *BIC* = Bayesian Information Criterion; *model with correlation of the errors of items 11 with 12 and 14 with 16.

To evaluate the reliability and validity of the scale, α, composite reliability and average variance extracted, and *AVE* were measured for each factor. The results can be seen in Table 4. The α values are acceptable and only the low engagement or irresponsibility dimension presented values < .70, although considering the stipulations of Taylor et al. (2008), given the low number of items, this value can be accepted. All factors present acceptable composite reliability values (Hair et al., 2014). In relation to the *AVE*, these same authors indicate that convergent validity values are considered acceptable when all the values of the standardized regression weights in a latent variable are significant and > .50, even if its *AVE* is < .50. This is the case for low engagement or irresponsibility and fails to follow directions. H2 is, thus, satisfied.

Table 3. Standardized factor loadings for first-order CFA and H-CFA of the CCD-EF

	CFA				*H-CFA*				
Item	*LEI*	*FFD*	*DDO*	*PSM*	*LEI*	*FFD*	*DDO*	*PSM*	*DB*
Low engagement or irresponsibility (LEI)									.842**
Item5	.570**				.578**				
Item6	.562**				.568**				
Item7	.662**				.653**				
Item8	.588**				.690**				
Fails to follow directions (FFD)									.972**
Item9		.761**				.762**			
Item10		.515**				.506**			
Item11		.525**				.528**			
Item12		.668**				.674**			
Distract or disturbs others (DDO)									.974**
Item13			.577**				.576**		
Item14			.745**				.748**		
Item15			.794**				.794**		
Item16			.743**				.743**		
Poor self-management (PSM)									.944**
Item17				.720**				.722**	
Item18				.696**				.697**	
Item19				.354**				.353*	
Item20				.713**				.711**	
Correlations									
FFD	.892**								
DDO	.792**	.938**							
PSM	.766**	.898**	.944**						

Note: *CFA* = first order confirmatory factorial analysis; *H-CFA* = higher order CFA; *DB* = disruptive behaviors; **$p < 0.01$.

Table 4. Reliability and validity of the CCD-EF factors

Factors	Composite Reliability	*AVE*	Cronbach´s Alpha
Low engagement or irresponsibility	.72	.38	.69
Fails to follow directions	.71	.39	.75
Distract or disturbs others	.81	.52	.81
Poor self-management	.72	.41	.70

Note: *AVE* = average variance extracted.

Factorial invariance across gender

The invariance of the CCD-EF was evaluated across gender (i.e., male = 375, female = 378) based on the first order CFA model and the higher order (H-CFA) model, the results of which are shown in Table 5. Starting with a configural invariance model, invariance constraints were progressively added to the factor loadings (i.e., weak invariance, intercepts (i.e., strong invariance), and residual variances (i.e., strict invariance), weak invariance), intercepts (i.e., strong invariance), and residual variances (i.e., strict invariance). The values of these restrictive models were acceptable, except for the strict invariance of the H-CFA, as the *CFI* results were outside the cut-off values. The values of these restrictive models of the H-CFA did not exceed the cut-off points for *RMSEA* ($\Delta > .015$), *CFI* ($\Delta > .01$), and *TLI* ($\Delta > .01$) so it can be considered fully invariant. In the case of the first-order model, it can be considered partially invariant, since the strict invariance model showed a decrease that slightly exceeded the limits of the recommended values (ΔCFI = -.013). H3 is, thus, fulfilled.

Table 5. Invariance test across gender for the CCD-EF

Model	χ^2	*df*	*RMSEA [90% IC]*	*CFI*	*TLI*	*ΔRMSEA*	*ΔCFI*	*ΔTLI*
Measurement across gender (First order CFA)								
1.- Configural invariance	489.542*	192	.045 [.040;.050]	.936	.920			
2.- Weak invariance	505.856*	204	.044 [.040;.049]	.935	.924	- .001	- .001	.004
3.- Strong invariance	553.927*	214	.046 [.041;.051]	.927	.918	.002	- .008	- .006
4.- Strict invariance	632.163*	232	.048 [.043;.052]	.914	.911	.002	- .013	- .007
Measurement across gender (H-CFA)								
1.- Configural invariance	514.016*	196	.046 [.042;.051]	.932	.917			
2.- Weak invariance	529.478*	208	.045 [.041;.050]	.931	.921	-.001	- .001	.004
3.- Strong invariance	554.529*	212	.046 [.042;.051]	.927	.917	.001	- .004	- .004
4.- Strict invariance	95.473*	212	.046 [.042;.051]	.927	.917	.000	.000	.000

Note: *CFA* = first order confirmatory factorial analysis; *H-CFA* = higher order *CFA*; *χ 2*= Chi square; *df* = degrees of freedom; *RMSEA* = root mean square error of approximation; 90% *CI* = 90% confidence interval of the *RMSEA*; *CFI* = comparative fit index; *TLI* = Tucker–Lewis index; * *p* < .01.

Nomological validity

A regression analysis with latent variables was performed to test the extent to which the dimensions of the SSI-EF (independent variable) predict disruptive behaviors (dependent variable). Firstly, to indicate that the SSI-EF showed excellent goodness-of-fit indices in CFA: *χ2* = 34.30; *gl* = 19; *p* = .01; *χ2/gl* = 1.80; *GFI* = .99; *CFI* = .97, *RMSEA* = .04, *SRMR* = .05, as well as adequate reliability: satisfaction/fun, *CR* = .84; *AVE* = .52; *α* = 78; boredom, *CR* = .71; *AVE* = .45; *α* = .65. Next, the measure of invariance of the linear regression model was tested according to the sex variable (girls, *n* = 378; boys, *n* = 375). The regression model was found to be invariant by sex, as the restrictive increase/decrease of the models for *RMSEA* was < .015, *CFI* < .01, and *TLI* < .01, the results are shown in Figure 1. In both boys and girls, the prediction of satisfaction/fun with PE on disruptive behaviors was not significant (H4 is not satisfied), but it was significant for boredom with PE (H5 is satisfied). When males are bored in PE classes, the predictive relationship for disruptive behaviors is .51 (*p* < .0001), with 23% of variance explained; while among female students the predictive relationship of boredom with PE for disruptive behaviors in class is .36 (*p* < .0001), with 15% of variance explained.

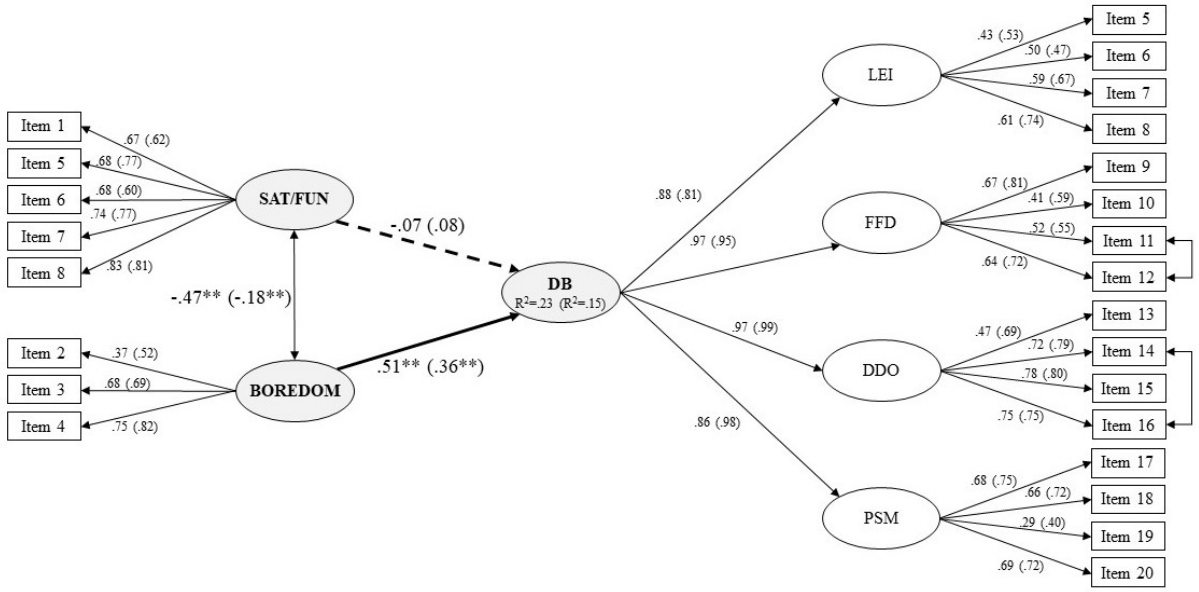

Figure 1. Structural regression model of satisfaction/fun and boredom with physical education classes and disruptive behaviors (higher order model)

Note: girls' values are in parentheses; *SAT/FUN*=satisfaction/fun; *DB*=disruptive behaviors; *LEI*=Low engagement or irresponsibility; *FFD*=Fails to follow directions; *DDO*=Distract or disturbs others; *PSM*=poor self-management. In parentheses are the values corresponding to the girls' model.

Discussion

The objectives of the present study were: (i) to analyze the psychometric properties of the CCD-EF scale in the Mexican context; (ii) to analyze a higher-order model based on the subscales that make up the CCD-EF; (iii) to analyze the factorial invariance of the CCD-EF scale as a function of sex; and (iv) to analyze the predictive relationship of satisfaction with PE on disruptive behaviors. The main results show that both the model with four correlated factors (first order) and the higher-order factor have adequate goodness-of-fit indices. In addition, the scale is invariant according to sex, and boredom with PE is a significant and positive predictor, stronger in males, of disruptive behaviors in the school context of secondary education in Mexico.

Concerning H1, the analysis of the data shows that the internal consistency obtained by the CCD-EF is adequate in all its dimensions, except for aggressive, which does not reach the minimum score to be accepted (Nunnaly & Bernstein, 1995). This same problem with the aggressive factor was already found in the validation of the instrument with Spanish adolescents (Granero-Gallegos & Baena-Extremera, 2016), it not being possible to make a comparison with the original English version, since it did not report internal consistency values. This result obtained in the aggressive dimension should be considered for the elimination of the factor. It is worth noting that the rest of the indicators suggest maintaining all the items of the scale (Nunnally & Bernstein, 1995).

Regarding the data obtained from the goodness-of-fit index, the present study analyzed several models following the indications of authors such as Markland (2007). The five-factor model did not reach the minimum acceptable values set by authors such as Hooper et al. (2008) to consider the CCD-EF as a valid instrument applicable in the Mexican context, so it was decided to eliminate the aggressive factor from the instrument since it presented problems in both reliability and validity analyses. Similar results were previously obtained by Granero-Gallegos and Baena-Extremera (2016) in their validation with a Spanish sample, in which they also proposed the elimination of the aggressive dimension. By contrast, in the original version, this factor was maintained since the results obtained in the validity of the instrument were acceptable. Subsequently, the instrument was analyzed with only four factors (i.e., low engagement or irresponsibility, fails to follow directions, distract or disturbs others and poor self-management), in which an excellent fit was obtained if two errors were correlated, one between items 11-12 and the other between items 14-16. Also, in the Spanish version (Granero-Gallegos & Baena-Extremera, 2016) an excellent fit was obtained in the four-factor model proposed. Thus, we can state that the four-dimensional CCD-EF is a valid and reliable instrument, as it has been used in different countries, and in the adaptation to the Mexican context, good results for reliability, validity and goodness-of-fit indexes have been obtained. In addition to this, it is worth highlighting the important contribution that the present study makes to the Mexican educational system and society, since it provides an instrument that evaluates the negative behaviors of adolescents in the PE classroom, in the country with the highest rate of violence among its elementary and high school students (Senado de la República de México, 2019). Therefore, H1 is fulfilled.

Regarding H2, the present work showed a higher-order factor of a reflective type (Lévy-Mangin & Varela, 2006) from the four subscales, obtaining adequate goodness-of-fit values. This means that the same student can have irresponsible behaviors, disobey the rules, be disruptive in class, and/or have no control over it, but, in addition, the CCD-EF can be used as a single latent variable that measures disruptive behaviors in general. This makes a contribution to the scientific literature that had not previously been made either in the original validation of the PECI (Krech et al., 2003; Krech et al., 2010) or in the adaptation and validation to the Spanish context carried out by Granero-Gallegos and Baena-Extremera (2016). Therefore, H2 is fulfilled.

In relation to H3, the present study has obtained that the Mexican version of the instrument is invariant according to sex, both for the four-factor model and for the higher order factor model. Thus, the factorial invariance of the CCD-EF is confirmed, and the CCD-EF can be considered to be valid and invariant instrument in the Mexican context for comparison studies based on the sex variable. Given its strength in the known psychometric properties of the CCD-EF (Granero-Gallegos & Baena-Extremera, 2016; Martínez-Molina et al., 2020), and the use in other languages (Krech et al., 2010), the data obtained have the potential to confirm construct validity. Thus, this study represents a contribution to the consistency of the psychometric properties of the CCD-EF, so that it can be used in future research regardless of gender. Therefore, H3 is confirmed.

In relation to H4 and H5, the prediction model proposed in this study showed that boredom with PE classes is a positive and statistically significant predictor of disruptive behaviors, both in girls and boys, although with higher predictive values in boys. This gender difference could be due to the fact that boys engage in negative behaviors more frequently than girls in the PE classroom (Baños, 2021; Cothran & Kulinna, 2007; Kulina et al., 2006). This can be explained by boys pretending to show superiority over their peers in order to attract the attention of girls (Glock & Kleen, 2017). Similar results were found other studies in which boredom in PE classes or at school predicted CCD-EF dimensions (Baños, 2020; Cothran et al., 2009; Granero-Gallegos et al., 2020). This could be due to the learning climates created by the teacher since when environments are created where the adolescent has to demonstrate superiority to the rest of his or her peers or avoid not showing incompetence (Hansen & Rindgal, 2018; Manzano-Sánchez et al., 2023) conflicts and disruptive behaviors among students are generated

(Baños, 2021). This highlights the importance of students not experiencing feelings of boredom as they will increase disruptive behaviors of some students in the classroom, who will then try to distract their peers in classes (Glock & Kleen, 2017). However, satisfaction and enjoyment with PE do not significantly predict CCD-EF, although it obtained a negative relationship. In this line Baños (2021) and Cothran et al. (2009), did obtain a statistically significant negative relationship between the dimensions of CCD-EF and satisfaction with PE, with the exception of the dimension of aggressive at work by Baños (2021), a factor that has been eliminated in the present study. It is worth mentioning that the scientific literature highlights the importance of students feeling satisfied at school as it reduces disruptive behaviors in both boys and girls (Asun-Dieste & Guíu, 2023; Aznar-Ballesta & Vernetta, 2023; Sun, 2016), increases satisfaction with life (Scharenberg, 2016) and is a relevant factor in the competencies acquired by the teacher for the creation of ideal learning climates for the student body (Invernizzi et al., 2019). Therefore, H5 is fulfilled, but not H4.

Conclusion

To conclude, the findings of this study have demonstrated that the CCD-EF is a reliable and valid instrument for application in the Mexican context and that it can be applied independently of the sex variable. Furthermore, it has been shown that the CCD-EF can be used as a higher order model, so that this scale can measure disruptive behaviors in a unidimensional way and, in turn, low engagement or irresponsibility, fails to follow directions, distract or disturbs others and poor self-management. This is a significant contribution to the existing literature since no scale of negative behaviors in PE had previously been validated with Mexican adolescents. In addition, boredom with PE positively and significantly predicts disruptive behaviors in both girls and boys.

Limitations and strengths

The present research has a series of strengths that should be mentioned, among them is the sample design, which is probabilistic and randomized by centers, stratified, multistage and with proportional allocation. In this way, the results of the study can be generalized to the State of Nuevo León in Mexico. Another important strength is the subject matter it addresses since it can contribute to generating solutions to the main problems related to adolescent behavior at school, a very worrying problem in Mexican society. It is also necessary to admit a series of limitations. For example, the CCD-EF scale can be applied at the high school level regardless of the subject's gender; however, it cannot be held that this scale can be used at different educational levels, so the results of the present study cannot be generalized to students of other ages. We consider it important to mention the need to carry out future studies of a longitudinal nature and using a mixed research methodology (quantitative and qualitative), also classifying the subjects at both socio-demographic and socio-economic levels.

Funding

This work was supported by the Secretariat of Public Education of Mexico (identification number: 431/569/E). This article was carried out related to a research stay of Dr. Raúl Baños at the University of Granada (from June 15, 2020, to August 11, 2020) with Dr. Antonio Baena Extremera.

References

Asun-Dieste, S., & Guíu Carrera, M. (2023). Use of formative assessment in physical education teacher education in secondary school: a case study. *Cultura, Ciencia y Deporte*, *18*(55), 171-190. https://doi.org/10.12800/ccd.v18i55.1918

Aznar-Ballesta, A., & Vernetta, M. (2023). Influencia de la satisfacción e importancia de la educación física en el abandono deportivo. *Espiral. Cuadernos del Profesorado*, *16*(32), 18-28. https://doi.org/10.25115/ecp.v16i32.8604

Baena-Extremera, A., Granero-Gallegos, A., & Martínez-Molina, M. (2015). Spanish version of the evaluation of teaching competencies scale in physical education of secondary school. *Cuadernos de Psicología del Deporte*, *15*(3), 113-122. https://dx.doi.org/10.4321/S1578-84232015000300011

Baños R. (2020). Boredom and disruptive behaviours in the physical education classrom. *Journal of Sport and Health Research, 12*(3), 406-419. https://recyt.fecyt.es/index.php/JSHR/article/view/83570

Baños R. (2021). Clima motivacional y conductas disruptivas en educación física en estudiantes españoles y mexicanos de educación secundaria. *Journal of Sport and Health Research, 13*(1), 1-12. https://recyt.fecyt.es/index.php/JSHR/article/view/87340

Baños, R., Baena-Extremera, A., Ortiz-Camacho, M. M., Zamarripa, J., Beltrán de la Fuente, A., & Juvera Portilla, J. L. (2019**)**. Influence of the competences of secondary teachers on disruptive behavior in the classroom. *Espiral. Cuadernos del Profesorado*, *12*(24), 3-10. https://doi.org/10.25115/ecp.v12i24.2141

Baños, R., Barretos-Ruvalcaba, M., & Baena-Extremera, A. (2019). Protocol for the study of the academic, psychological and physical activity variables that influence the academic performance of Mexican and Spanish adolescents. *Espiral. Cuadernos del Profesorado*, *12*(25), 89–99. https://doi.org/10.25115/ecp.v12i25.2480

Baños**,** R., Baena-Extremera, A., & Granero-Gallegos**,** A. (2019). The relationships between high school subjects in terms of school satisfaction and academic performance

in Mexican adolescents. *International Journal of Environmental Research and Public Health*, *16*(18), 3494. https://doi.org/10.3390/ijerph16183494

Baños, R., Ortiz-Camacho, M. M., Baena-Extremera, A., & Tristán-Rodríguez, J. L. (2017). Satisfaction, motivation and academic performance in students of secondary and high school: background, design, methodology and proposal of analysis for a research paper. *Espiral. Cuadernos del Profesorado*, *10*(20), 40–50. https://doi.org/10.25115/ecp.v10i20.1011

Benbenishty. R., Astor, R. A., & Roziner, I. (2018). A school-based multilevel study of adolescent suicide ideation in California high schools. *The Journal of Pediatrics*, *196*, 251-257. https://doi.org/10.1016/j.jpeds.2017.12.070

Bentler, P.M. (2007). On tests and indices for evaluating structural models. *Personality and Individual Differences, 42*(5), 825–829. https://doi.org/10.1016/j.paid.2006.09.024

Burns, T., & Gottschalk F. (2019). *Educating 21st Century children: emotional well-being in the digital age. educational research and innovation*. OECD**.** https://www.oecd.org/education/educating-21st-century-children-b7f33425-en.htm

Buscá, F., Ruiz, L., & Rekalde, I. (2014). Conflict resolution in schools as Learning Communities through physical education. *Retos. Nuevas tendencias en Educación Física, Deporte y Recreación*, 25, 156-161. https://doi.org/10.47197/retos.v0i25.34503

Byrne, B.M. (2013). *Structural equation modeling with AMOS: basic concepts, applications, and programming*. Routledge.

Carretero-Dios, H., & Pérez, C. (2007). Standards for the development and the review of instrumental studies: considerations about test selection in psychological research. *International Journal of Clinical and Health Psychology, 7*, 863-882. https://www.aepc.es/ijchp/carretero_es.pdf

Castillo, I., Balaguer, I., & Duda, J. L. (2001). Las perspectivas de meta de los adolescentes en el contexto deportivo. *Psicothema, 14*(2), 280–287. https://www.psicothema.com/pdf/721.pdf

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1)**,** 155-159. https://doi.org/10.1037//0033-2909.112.1.155

Cothran, D. J., Kulinna, P. H., & Garrahy, D. (2009). Attributions for and consequences of student misbehavior. *Physical Education and Sport Pedagogy, 14*(2), 155-167. https://doi.org/10.1080/17408980701712148

Cothran, D. J., & Kulinna, P. H. (2007). Students' reports of misbehavior in physical education. *Research Quaterly for Exercise and Sport, 78*(3)**,** 216 -224. https://doi.org/10.1080/02701367.2007.10599419

De Agüero, M. (2020). Research on bullying and violence at school in Mexico. *Revista Digital Universitaria, 21*(4). http://doi.org/10.22201/cuaieed.16076079e.2020.21.4.2

Declaration of Helsinki (2013). *Ethical principles for medical research involving humans*. World Medical Association. https://www.wma.net/what-we-do/education/medical-ethics-manual/

Domínguez-Lara, S. (2017). Effect size in regression analysis. *Interacciones. Journal of Family, Clinical and Health Psychology*, *3*(1), 3-5. https://doi.org/10.24016/2017.v3n1.46

Fernández-Garcimartín, C., López Pastor, V. M., Fuentes-Nieto, T., & Hortigüela-Alcalá, D. (2023). Formative and shared assessment and final degree projects in physical education pre-service teacher education. *Cultura, Ciencia y Deporte*, *18*(55), 33-56. https://doi.org/10.12800/ccd.v18i55.1922

Glock, S., & Kleen, H. (2017). Gender and student misbehavior: evidence from implicit and explicit measures. *Teaching and Teacher Education 67*, 93–103. https://doi.org/10.1016/j.tate.2017.05.015

Granero-Gallegos, A., & Baena-Extremera, A. (2016). Validation of the short-form Spanish version of the physical education classroom instrument measuring secondary pupils' disruptive behaviours. *Cuadernos de Psicología del Deporte*, *16*(2), 89-98. https://revistas.um.es/cpd/article/view/264451

Granero-Gallegos, A., Baños, R., Baena-Extremera, A., & Martínez-Molina, M. (2020). Analysis of misbehaviours and satisfaction with school in secondary education according to student gender and teaching competence. *Frontiers in Psychology*, *11*(63). https://doi.org/10.3389/fpsyg.2020.00063

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Prentice Hall.

Hansen, G., & Ringdal, R. (2018). Formative assessment as a future step in maintaining the mastery-approach and performance-avoidance goal stability. *Studies in Educational Evaluation, 56*, 59-70. https://doi.org/10.1016/j.stueduc.2017.11.005

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115-135. https://doi.org/10.1007/s11747-014-0403-8

Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodología de la investigación* **(6a Ed.).** McGraw-Hill.

Hinduja, S., & Patchin, J. W. (2015). *Bullying beyond the schoolyard: preventing and responding to cyberbullying* (2nd ed.). Sage Publications.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53-60. https://doi.org/10.21427/D7CF7R

Hoyo-Guillot, A., & Ruiz-Montero, P. J. (2023). Influence of a personal and social responsibility model intervention on educative motivation, prosocial environment and physical education importance on secondary school students. *Espiral. Cuadernos del Profesorado*, *16*(34), 53-70. https://doi.org/10.25115/ecp.v16i34.9629

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria. *Structural Equation Modeling: a Multidisciplinary Journal, 6*(1): 1-55. https://doi.org/10.1080/10705519909540118

Invernizzi, P. L., Crotti, M., Bosio, A., Cavaggioni, L., Alberti., G., & Scurati, R. (2019). Multi-teaching styles approach and active reflection: effectiveness in improving fitness level, motor competence, enjoyment, amount of physical activity, and effects on the perception of physical education lessons in primary school children. *Sustainability 11*(2), 405**.** https://doi.org/10.3390/su11020405

Krech, P. R., Kulinna, P. H., & Cothran, D. (2010). Development of a short-form version of the physical education classroom instrument: measuring secondary pupils' disruptive behaviours. *Physical Education and Sport Pedagogy, 15*(3), 209–225. https://doi.org/10.1080/17408980903150121

Kulinna, P. H., Cothran, D., & Regualos, R. (2003). Development of an instrument to measure student disruptive behavior. *Measurement in Physical Education and Exercise Science*, *7*(1), 25-41. https://asu.elsevierpure.com/en/publications/development-of-an-instrument-to-measure-student-disruptive-behavi

Kulinna, P. H., Cothran, D., & Regualos, R. (2006). Teachers' reports of student misbehavior in physical education. *Research Quarterly for Exercise and Sport 77*(1), 32–40. https://doi.org/10.1080/02701367.2006.10599329

Lévy-Mangin, J. P., & Varela, J. (2006). *Modelización con estructuras de covarianzas en ciencias sociales*. Netbiblo.

Lynn, M. (1986). Determination and quantification of content validity. *Nursing Research*, *35*(6), 382-385. https://pubmed.ncbi.nlm.nih.gov/3640358/

Maddeh, T., Bennour, N., & Souissi, N. (2015). Study of students' disruptive behavior in high school education in physical education classes. *Advances in Physical Education, 5*(3), 143-151. https://doi.org/10.4236/ape.2015.53018

Manzano, D. (2021). Diferencias entre aspectos psicológicos en Educación Primaria y Educación Secundaria: motivación, necesidades psicológicas básicas, responsabilidad, clima de aula, conductas antisociales y violencia. *Espiral. Cuadernos del Profesorado, 14*(28), 9-18. https://doi.org/10.25115/ecp.v14i28.4198

Manzano-Sánchez, D., Ferrer-López, B., Battaglia, G., & Gómez-López, M. (2023). Support for autonomy and fear of failure in physical education lessons. Differences between "gender and physical extracurricular sports. *Espiral. Cuadernos del Profesorado*, *16*(34), 1-11. https://doi.org/10.25115/ecp.v16i34.9476

Markland, D. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modelling. *Personality and Individual Differences*, *42*(5), 851-858. https://doi.org/10.1016/j.paid.2006.09.023

Marsh, H. W., Hau, K. T., & Wen, Z. L. (2004). In search of golden rules: comment on approaches to setting cutoff values for fit indexes and dangers in over generalising Hu & Bentler (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal, 11*(3), 320-341. https://doi.org/10.1207/s15328007sem1103_2

Martínez-Molina, M., Granero-Gallegos, A., Baena-Extremera, A., & Baños, R. (2020). Factorial invariance by sex of the physical education classroom instrument and evaluation of teaching competencies scale adapted to physical education with secondary students. *Psychology, Society, & Education*, *12*(2), 125-137. https://repositorio.ual.es/bitstream/handle/10835/8671/3172-10639-1-PB.pdf?sequence=1&isAllowed=y

Milfont, T. L., & Fisher, R. (2010). Testing measurement invariance across groups: applications in cross-cultural research. *International Journal of Psychological Research*, *3*(1), 111–121. https://revistas.usb.edu.co/index.php/IJPR/article/view/857/607

Myers, N. D., Martin, J. J., Ntoumanis, N., Celimli, S., & Bartholomew, K. (2014). Exploratory bifactor analysis in sport, exercise, and performance psychology: a substantive-methodological synergy. *Sport, Exercise, and Performance Psychology*, *3*(4), 258-272. https://doi.org/10.1037/spy0000015

Nunnally, J.C., & Bernstein, I.J. (1995). *Teoría psicométrica*. Mcgraw-Hill.

Olweus, D., & Breivik, K. (2014). Plight of victims of school bullying: the opposite of well-being. In A. Ben-Arieh, F. Casas, I. Frønes & J. E. Korbin (Eds.), *Handbook of Child Well-Being SE - 90* (pp. 2593–2616). Springer.

Palacios-Gómez, J., Hortigüela-Alcalá, D., & Hernando-Garijo, A. (2023). Fears and perceptions of the application of formative and shared assessment by a

novel physical education teacher. *Espiral. Cuadernos del Profesorado*, *16*(34), 28-39. https://doi.org/10.25115/ecp.v16i34.9627

Rindskopf, D., & Rose, T. (1988). Some theory and applications of confirmatory second- order factor analysis. *Multivariate Behavioral Research*, 23, 51-67. https://doi.org/10.1207/s15327906mbr2301_3

Scharenberg, K. (2016). The interplay of social and ethnic classroom composition, tracking, and gender on students' school satisfaction. *Journal of Cognitive Education and Psychology. 15*(2), 320–346. https://doi.org/10.1891/1945-8959.15.2.320

Senado de la República de México (2019). Boletín: México, en los primeros lugares del mundo en acoso escolar. Publicado miércoles, 21 de agosto del 2019.

SEP (2017). *Aprendizajes clave para la formación integral*. SEP. https://www.gob.mx/sep/articulos/aprendizajes-clave-para-la-educacion-integral

Soper, D. S. (2022). *A-priori sample size calculator for Student t-Tests.* Free Statistics Calculators. https://www.danielsoper.com/statcalc/

Sulbarán, A., & León, A. (2014). Estudio de las conductas disruptivas en la escuela según la percepción docente. *Administración Educacional, 2*, 35-50. http://erevistas.saber.ula.ve/index.php/administracioneducacional/article/view/5302

Sun, R.C. (2016). Student misbehavior in Hong Kong: the predictive role of positive youth development and school satisfaction. *Applied Research in Quality of Life 11*, 773–789. https://doi.org/10.1007/s11482-015-9395-x

Supaporn, S., Dodds, P., & Griffin, L. (2003). An ecological analysis of middle school misbehavior through student and teacher perspectives. *Journal of Teaching in Physical Education*, *22*(3), 328-349. https://doi.org/10.1123/jtpe.22.3.328

Tabachnick, B.G., & Fidell, L. S. (2018). *Using Multivariate Statistics*. Pearson.

Taylor, I., Ntoumanis, N., & Standage, M. (2008). A self-determination theory approach to understanding antecedents of teachers' motivational strategies in physical education. *Journal of Sport and Exercise Psychology*, *30*(1), 75-94. https://doi.org/10.1123/jsep.30.1.75

Teo, T. S. H., Srivastava, S.C., & Jiang, L. (2008). Trust and electronic government success: an empirical study. *Journal of Management Information Systems*, *25*(3), 99-132. https://doi.org/10.2753/MIS0742-1222250303

Von-Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P. (2008). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Bulletin of the World Health Organization, 61*(4), 344-349. https://doi.org/10.1016/j.jclinepi.2007.11.008