Espiral. Cuadernos del Profesorado
Vol. 16, Num. 33
September 2023
ISSN 1988-7701

How to reference this article:

Baños, R., Machado-Parra, J.P., Arrayales-Millán, E.M., & Baena-Extremera, A. (2023). Psychometric properties of the Teaching Competence Assessment Scale in Mathematics. Satisfaction and academic performance prediction model. *Espiral. Cuadernos del Profesorado, 16*(33), 1-13. https://doi.org/10.25115/ecp.v16i33.9369

Psychometric properties of the Teaching Competence Assessment Scale in Mathematics. Satisfaction and academic performance prediction model

Propiedades psicométricas de la Escala de Evaluación de la Competencia Docente en Matemáticas. Modelo de predicción de satisfacción y rendimiento académico

Raúl Baños[1], Juan Pablo Machado-Parra[1], Emilio Manuel Arrayales-Millán[1], and Antonio Baena-Extremera[2]

[1]Autonomous University of Baja California, México

[2]University of Granada, Spain

Abstract

The purposes of this study were: (i) to provide reliability and validity evidence regarding the Teaching Competence Questionnaire for the Math class in the Mexican context; (ii) to analyze factorial invariance according to sex in the Teaching Competence Questionnaire; (iii) to study the relationship among student perceptions of Math teaching competencies, satisfaction with school, and academic performance. A sample was selected from the total number of third-grade secondary school Mexican students. The prediction model showed that the students' perception of their Math teachers' competence does predict satisfaction with Math class, as well as their academic performance in Math.

Keywords: teaching competencies; validation; Math; satisfaction; academic performance.

Resumen

Los objetivos de este estudio fueron: (i) proporcionar evidencia de confiabilidad y validez del Cuestionario de Competencias Docentes para la clase de Matemáticas en el contexto mexicano; (ii) analizar la invariancia factorial según sexo en el Cuestionario de Competencia Docente; (iii) estudiar la relación entre las percepciones de los estudiantes sobre las competencias de enseñanza de Matemáticas, la satisfacción con la escuela y el rendimiento académico. Se seleccionó una muestra del total de estudiantes mexicanos de tercer grado de secundaria. El modelo de predicción mostró que la percepción de los estudiantes sobre la competencia de sus profesores de Matemáticas sí predice la satisfacción con la clase de Matemáticas, así como su rendimiento académico en Matemáticas.

Palabras clave: competencias docentes; validación; matemáticas; satisfacción; rendimiento académico.

Received: July 1, 2023
Acepted: July 17 2023

Correspondence: Raúl Baños, Autonomous University of Baja California, Mexico
Email: raulfb89@gmail.com

Introduction

Governments worldwide and national ministries of education have focused their attention on improving their educational systems. The countless factors involved in the teaching-learning process of an educational system may increase or decrease the probability of its success depending on the interaction between them (Baños et al., 2019a). Among these factors, the role of teachers is of great relevance, especially when considering the results attained in international external reports. Hence, the competencies of teachers and the assessment of their professional career are gradually becoming questions in need of being addressed (Baena-Extremera et al., 2015).

At the international level, the Organization for Economic Co-operation and Development (OECD) has carried out the Teaching and Learning International Study (TALIS), known for assessing the most effective teaching competencies. As a consequence of working in more challenging environments, feeling less prepared to do their job, and scoring lower in academic training, Mexican teachers have obtained concerning results (OCDE, 2013), which, in turn, might have triggered the low scores obtained by Mexican adolescents in the PISA Report at a general level, and specifically for the Math knowledge assessment (MAT), scoring 409 points, whereas the average for OECD countries was 489 (OCDE, 2019). The Secretariat of Public Education of Mexico (SEP) applies the National Plan for Learnings Assessment (PLANEA) nationwide, whose 2017 report showed an average score decline compared to the 2015 baseline; in terms of language and communication, merely 47% of students achieve the minimum essential threshold, whereas results in MAT are even more unsettling, with two-thirds of students not acquiring the expected learnings (PLANEA, 2017). When searching for potential causes, it is believed from previous research efforts that academic performance and the behavior of students in the classroom are related to the teachers' self-perception of their own abilities, as shown by the international literature, which is perhaps not very promising (Espeland & Indrehus 2003; Lizzio et al., 2002; Paisey et al., 2007). However, no studies have compared these variables in Mexico nor analyzed the perception that Mexican students have of their teachers' competencies.

In this line, the quality of education depends largely on the quality of teachers; apart from possessing skills in teaching and researching, they must also prove to be academic leaders. However, some authors see the lack of training in pedagogical skills as the main reason to why teachers are hindering students' learnings (Fawad & Manarvi, 2014). In fact, researchers have underscored the need to drive changes in teaching methods, particularly with how theoretical classes are taught (Baños et al., 2021). One of the reasons is that students feel that teachers merely provide learning content without developing other skills in them (Hodgson et al., 2014). On the contrary, when learning topics that awaken their interest, students feel heard by their teachers, they acquire new information, they sense a spirited environment and that assigned tasks are based on projects, and they improve their learning during class (Shomoossi 2004; Weber et al., 2001). As can be seen, the literature shows two sides of the teaching practice and how these may impact the teaching and learning process. Having explained this though, what parties are the most adequate to assess teaching competencies?

Several authors underscore the importance of students as assessment agents to maintain the quality of learning and improve learning by participating in the evaluation processes that guarantee the internal quality of teaching (Lidice & Saglam, 2012), and some questionnaires meant to measure self-perception of acquired learnings and the student's academic experience have been based on this idea (Benton et al., 2011; Fawad & Manarvi, 2014). In this sense, Douglas, Thomson and Zhao (2012), found that the student's self-perception of their specialization or learning program level was the best indicator for assessing their own academic experience. The existing relationship between the students' self-evaluations and better academic performance is also worth mentioning (McNaught et al., 2012; Benton et al., 2011). In addition, learning self-reports have been correlated with the grade point average and performance in standardized evaluations (Anaya, 1999). This is why several authors underline the importance of analyzing the assessment made by the students on their teachers' competencies (Baños et al., 2019a; Baena-Extremera et al., 2015). The main inconvenience found by the scientific literature is the lack of measurement

instruments that meet the reliability and validity criteria to assess teachers from the students' standpoint (Catano & Harvey, 2011; Marsh, 2007; Simpson & Siguaw, 2000).

The Teaching Competence Assessment Questionnaire (Cuestionario de Evaluación de la Competencia Docente) by Luna, Calderón, Caso and Cordero (2012) applied on university students is one of the measurement instruments found in literature designed to assess teaching competencies. This scale measures four factors: planning and management of the teaching-learning process, didactic interaction in the classroom, assessment and disclosure of the teaching-learning process, and information and communication technologies. Nonetheless, although the instrument designed by Luna et al. (2012) did obtain an acceptable internal consistency, its corresponding analyses were not reported, so it is unknown whether it meets the minimum requirements for instrument validity.

Another instrument developed for this purpose is the Evaluation of Teaching Competencies Scale (ETCS), validated by Catano and Harvey (2011) on university students and obtaining good reliability results. This instrument measures the requirements that a teacher must meet to be deemed competent (good communication, work awareness, creativity, feedback, individual consideration, professionalism, problem solving, and social awareness), however, the validity of the instrument was not analyzed either. The ETCS was later adapted for Physical Education and validated in Spanish with secondary school students, and renamed Questionnaire of Teaching Competencies (Cuestionario de Competencias Docentes, CCD; Baena-Extremera et al., 2015), yielding excellent results in terms of validity and reliability. By doing so, Baena-Extremera et al. (2015) contributed data on the goodness-of-fit index in terms of the construct validity, which neither Catano and Harvey (2011) nor Luna et al. (2012) had previously analyzed. Furthermore, it is worth mentioning that the CCD has been used in several studies with Spanish students (Baños et al., 2019a; Baños et al., 2018; Granero-Gallegos et al., 2020), demonstrating its substantial reliability and validity. Nevertheless, Baena-Extremera et al. (2015) did not analyze factorial invariance, an issue that Martínez-Molina et al. (2020) remediated recently, obtaining outstanding results. Henceforth, we consider the CCD to be an adequate instrument to be adapted for the MAT class and analyze the psychometric properties of a sample with Mexican secondary school students. In this way, we would obtain a cutting-edge, groundbreaking and reliable instrument capable of assessing the students' perception of their MAT teachers' competencies, while attempting to offer a possible explanation to the low scores obtained in the assessments of the PISA report, PLANEA, and the TALIS Project.

Scientific literature has shown that when a student believes that their teachers are competent, their own satisfaction with school improves (Invernizzi et al., 2019), while negative behaviors in the classroom decrease (Baños et al., 2019b). On the contrary, when students think that their teachers are incompetent, there is an increase in dissatisfaction with school (Baños et al., 2019a; Sun, 2016), and even potentially negative behaviors in the classroom (Rasmussen et al., 2014) and school dropout rates (Takakura et al., 2010). It has been recently demonstrated that dissatisfaction with the MAT class predicts boredom at school and low academic performance in Mexican adolescents, whereas satisfaction with the MAT class predicts satisfaction with school and good academic performance (Baños et al., 2019c). Furthermore, the students' perception of the environment created by their teacher in the MAT classroom predicts their subject scores (Marsh et al., 2008; Samuelsson & Samuelsson, 2016). As mentioned above, both the students' satisfaction and the importance of perceiving their MAT teachers as competent are relevant for their MAT class scores. However, we did not find any studies that have linked the students' perception of their MAT teachers' competencies to satisfaction and academic performance in the MAT class. This study was devised based on that idea.

It is worth underscoring the importance of assessing student perceptions of their MAT teachers' competencies according to sex. Several studies have found differences in variables such as achievement goals, perception of the classroom environment, and academic performance in MAT class, with girls obtaining higher scores (Gherasim et al., 2013). Girls also reported receiving more support from their MAT teacher in the classroom (Shim et al., 2008) and better understanding the learning contents (Pekrun et al., 2006). Nonetheless, both boys and girls underscored the importance of their MAT teacher's skills and behavior in terms of their learning process, particularly stressing that they should be receptive, helpful and supportive (Ahmed et al., 2010; Patrick et al., 2007; Puklek-Levpuscek, & Zupancic, 2009). Therefore,

we deemed it important to analyze the factorial invariance of the CCD regarding the MAT class, and to discern whether it is applicable for both sexs.

In view of the foregoing, we propose the following objectives: (i) to provide CCD reliability and validity evidence in the MAT class for the Mexican context; (ii) to analyze the factorial invariance of the CCD according to sex; (iii) to study the relationship among the students' perception of their MAT teacher's competencies, satisfaction with school, and academic performance. The following hypotheses are established: First, it is hypothesized that the CCD instrument will obtain adequate validity and reliability values in a sample of Mexican secondary school students (H1). Second, it is hypothesized that factorial invariance will yield adequate values according to sex (H2). Third, it is hypothesized that a student perceiving their MAT teachers as competent predicts satisfaction with school and high academic performance (H3), whereas perceptions of incompetence predict dissatisfaction with school and low academic performance (H4).

Method

Design and Procedure

The Physical Education CCD items by Baena-Extremera et al. (2015) underwent a back-translation process (Muñiz et al., 2013) for their adaptation into the Mexican context. Two native translators adapted the nine items into the Mexican context and, subsequently, two different translators rendered the items back into their original language (back-translation). The degree of similarity with the original version was the baseline to assess the accuracy of the translation. Four experts in secondary school teaching analyzed the final version based on Lynn (1986) to ensure an adequate design of the items according to the construct intended to be measured while preserving their original meaning. The experts assessed the appropriateness and comprehensibility of each item on a scale ranging from 1 (Completely disagree) to 4 (Completely agree) and revised any items with average scores <2.5 and/or those not classified within the theoretical dimensions of the scale by at least three of the four experts. The global agreement of the four experts on appropriateness and comprehensibility was measured using the Interclass Correlation Coefficient (CCI); values obtained were: CCI = 0.81 for appropriateness and CCI = 0.83 for comprehensibility. The Spanish version was applied to 45 secondary school students aged between 12 and 14 years old, who expressed a complete understanding of the items. The Mexican version of the CCD-MAT was thus created.

In order to conduct this study, a research project called "Programme for International Student Assessment: relationship between school performance in secondary school students and psychological, family and physical activity variables" was first presented to, and later approved and subsidized by the Secretariat of Public Education of Mexico (identification number: 431/569/E). Then, authorization was requested from secondary school principals, providing the parents/guardians involved with information for consent detailing the purpose and intentionality of the study. Following their approval, the data collection procedure began by informing the participants of the study's purpose, that participation was anonymous and voluntary, and that their answers were to remain confidential, reminding them that there were no right or wrong answers, and asking them to answer with complete honesty. Students filled all questionnaires inside the classroom in the presence of the lead researcher in case of doubts during the procedure, which lasted 15–20 minutes.

Participants

The sample design was probabilistic by centers, stratified, multistage and by proportional affixation, and was comprised of third-grade secondary school students from the State of Nuevo León (Mexico). Participant secondary schools were selected at random using a random number table. The total number of third-grade secondary school students in the State of Nuevo León was 13,396 girls and 13,831 boys. The representative sample indicated was calculated according to sex for a finite population with a confidence level of 95% and a margin of error of +5%, consisting of 374 girls (M_{age} = 13.99; *SD* = 0.30) and 374 boys (M_{age} = 14.02; *SD* = 0.33).

Instruments

The questionnaire comprehended the following scales:

The CCD instrument for Physical Education by Baena-Extremera et al. (2015) was adapted into the Mexican context and the MAT class to measure what perception students have of their MAT teachers' competencies (CCD-MAT). It comprises nine items that measure the students' perception of their teachers' efficiency. The scale was preceded by the phrase: "Tell us how much do you agree or disagree with your Math teacher's competence" (Dinos tu grado de acuerdo o desacuerdo en cuanto a la competencia de tu profesor de Matemáticas). Responses were given on a Likert scale for low (1, 2), medium (3, 4, 5) and high (6, 7).

The MAT Intrinsic Satisfaction instrument adapted into the Mexican context by Baños et al. (2019c) was used to measure intrinsic satisfaction during the MAT class (SIMAT) through the enjoyment and boredom dimensions. The instrument is comprised of eight items, five of which measure the degree of satisfaction/enjoyment with the academic activities of each class, while the remaining three measure dissatisfaction/boredom. The scale was preceded by the phrase: "Tell us how much do you agree or disagree with the MAT class" (Dinos tu grado de acuerdo o desacuerdo en cuanto a las clases de MAT). Answers were given on a Likert scale ranging from 1 (completely disagree) to 5 (completely agree). The alpha value found in the study by Baños et al., (2019c) for the satisfaction/enjoyment subscale was 0.75, and 0.65 for the boredom subscale. Fitness indices found for the confirmatory factor analysis were as follows: $\chi^2 = 20.41$; gl = 19; p = 0.37; χ^2/gl = 1.07; GFI = 0.99; NFI = 0.97; NNFI = 0.99; CFI = 0.99 and RMSEA = 0.01.

Lastly, we asked teachers for permission to access their student scores records. These were registered on a scale of polytomous items ranging from 1 to 10.

Statistical Analysis

As a first step to fulfill the objectives of this study, homogeneity statistical analyses for each item and the internal consistency study were performed using the SPSS 25.0 software package. The analysis of each of the CCD scale items followed the guidelines established by Carretero-Dios and Pérez (2007). This analysis is necessary to assess the convenience of keeping each item for measuring the theoretical construct according to the original scale (Catano & Harvey, 2011). The study of the items included analyzing the internal consistency of the scale amended with the elimination of an item, as well as the requirements established by Nunnally and Bernstein (1995) to preserve an item within a factor: corrected item-total correlation coefficient (CCIT-c) ≥0.30, standard deviation (SD) >1.0 and all answer options used at any given moment. As recommended by Bollen and Long (1994), skewness and kurtosis values approach 0 and <2.

The factorial structure of the analyzed instrument was assessed through confirmatory factor analysis (CFA) using the statistical software AMOS 24 (Arbuckle, 2015). Several goodness-of-fit indices were calculated to assess the models, combining both absolute and relative indices. Among absolute indices, the p-value was used associated to the chi-square (χ2) statistical test on the null model against the hypothesized model (Barret, 2007); the ratio between χ2 and degrees of freedom (gl) (χ2/gl), considering <2.0 as indicators for very good model fitness (Tabachnik & Fidell, 2007). The GFI (goodness-of-fit index) was calculated considering values ≥0.95 for better fit according to authors such as Hooper et al. (2008). In terms of relative indices, the NFI (normed fit index), NNFI (non-normed fit index) and CFI (comparative fit index) were calculated; values ≥0.95 indicate an adequate fit (Hu & Bentler, 1999). These authors consider that a value of ≤0.06 for the RMSEA (root mean square error of approximation) indicates a good fit, as well as ≤0.08 values for the RMSR (standardized root mean square residual). Estimated parameters are deemed significant when the value is associated with the t-value >1.96 (p<0.05).

Cronbach's alpha was subsequently used to assess the internal consistency of each dimension. In addition, it is also interesting to offer composite reliability and average variance extracted (AVE) data in the correlation matrix scale CFA.

The factorial invariance analyses were performed using AMOS 24. Invariance according to sex (multigroup invariance) was tested according to the advanced methodological proposal by Milfont and

Fisher (2010) of testing four increasingly constrained models; thus, four increasingly restricted nested models were considered: (1) Model 1, configural invariance (M1, no restrictions), which served as a benchmark for the other models, (2) Model 2, metric invariance (M2, restriction on factorial loads); Model 3, strong invariance (M3, restriction on factorial loads and intercepts); Model 4, strict invariance (M4, restriction on factorial loads, intercepts and error variance). Based on Chen (2007), said nested models were compared considering the change (Δ) in the goodness-of-fit indices (i.e., an RMSEA increase of at least 0.015 or a CFI and TLI decrease of at least 0.010 indicate a lack of invariance).

Lastly, the concurrent validity of the CCD-MAT was examined using AMOS 24 (Arbuckle, 2015). Two prediction models were proposed for its relationship with satisfaction/enjoyment and dissatisfaction/boredom at school using a prediction model through structural regression analysis.

Results

Descriptive Analyses of the Items and Exploratory Factor Analysis

Results from the descriptive analyses of the CCD-MAT items showed mean values ranging between 5.18 (item-1) and 5.68 (item-3). SD were > 1. Internal consistency was adequate (α = 0.90) and does not improve by removing any item. All CCIT-c had values ≥0.57. As can be seen in Table 1, results meet the values outlined in the literature for each of the descriptive statistics (Bollen & Long, 1994; Carretero-Dios & Pérez, 2007; Nunnally & Bernstein (1995).

The CCD-MAT instrument yielded significant results for the Bartlett sphericity test (χ 2 = 3054.045, gl = 36; p < 0.001), whereas Kaiser-Meyer-Olkin was above 0.60 (KMO = 0.94), signaling data adequacy. Results supported the structure of a factor with eigenvalues above 1, a total accumulated variance of 54.88% and a factorial load >0.40.

Table 1

Descriptive statistics for all items and exploratory factor analysis of the scale CCD-MAT (n = 748)

Items	*M*	*SD*	*CCIT-c*	α without item	Skewness	Kurtosis	λ
Item 1	5.18	1.80	.57	.89	-0.93	-.14	.65
Item 2	5.60	1.64	.67	.88	-1.14	0.41	.75
Item 3	5.68	1.74	.58	.89	-1.49	1.24	.66
Item 4	5.22	1.84	.79	.89	-0.90	-.27	.67
Item 5	5.51	1.75	.73	.88	-1.19	0.50	.80
Item 6	5.23	1.82	.71	.88	-1.01	0.05	.79
Item 7	5.35	1.79	.71	.88	-1.05	0.12	.78
Item 8	5.26	1.82	.74	.88	-0.98	-0.01	.81
Item 9	5.30	1.75	.65	.89	-1.05	0.21	.73

Note: *n* = sample; *M* = mean; *SD* = standard deviation; CCIT-c = corrected coefficient of item-total correlation; λ = Standardized factorial loads.

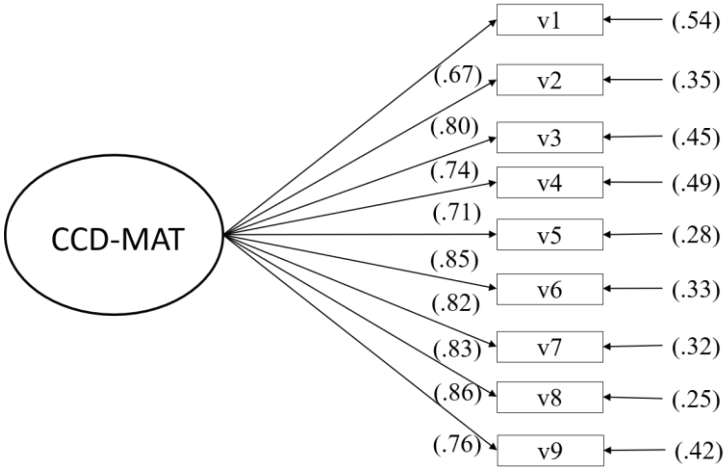
Confirmatory Factor Analysis

Structural equations modeling was applied in order to study the psychometric properties of the original dimensioning theoretically proposed by Catano and Harvey (2011). The factorial structure of the CCD-MAT was assessed by conducting a CFA.

The tested model possesses high factorial loads (≥ 0.67). All items showed values > 0.50 for individual reliability (R2) (ranging between 0.67 for item-1 and 0.86 for item-8). All fitness data were excellent: χ^2=54.40, gl=27, χ^2/gl=2.01; p=0.001, χ^2/gl=2.014, GFI=0.994, NFI=0.973, NNFI=0.982, CFI=0.986, RMSEA=0.037. The model meets the requirements to ensure its convergent validity (Hair et al., 2009): all factorial loads are >0.60 (see Figure 1) and statistically significant (t_{value}>1.96).

Figure 1

Path diagram de la escala CCD-MAT



Regarding the data correlation matrix ordinal scales CFA, it is also important to offer results for composite reliability and average variance extracted (AVE) for each of the critical dimensions. According to Hair et al. (2009), the minimum values for composite reliability and AVE should be 0.70 and 0.50, respectively. The tested model had a composite reliability of 0.94 and an AVE of 0.62

Factorial Invariance According to Sex

Invariance according to sex (i.e., 374 girls, 374 boys) of the CCD-MAT was based on CFA models. Invariance test results are shown in Table 2. All models (configural invariance, metric invariance, strong invariance, and strict invariance) provide a good goodness-of-fit level (CFI >0.90, TLI >0.90, and RMSEA<0.08) and do not exceed the cutoff points recommended for comparison of the increasingly restrictive models for RMSEA (Δ<0.015), CFI (Δ <0.01), and TLI (Δ<0.01).

Table 2

Goodness of fit indexes of the invariance models

Model	χ^2	*df*	RMSEA[90% IC]	CFI	TLI	Model comparison	ΔRMSEA	ΔCFI	ΔTLI
M1	163.977	54	.052 [.043-.061]	.965	.953	-	-	-	-
M2	181.161	62	.051 [.042-.059]	.962	.956	2 *versus* 1	.001	.003	-.003
M3	198.699	71	.049 [.041-.057]	.959	.958	3 *versus* 2	.002	.003	-.002
M4	220.101	81	.048 [.040-.055]	.955	.960	4 *versus* 3	.001	.004	-.002

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; * $p < .01$.

Structural Equations Modeling

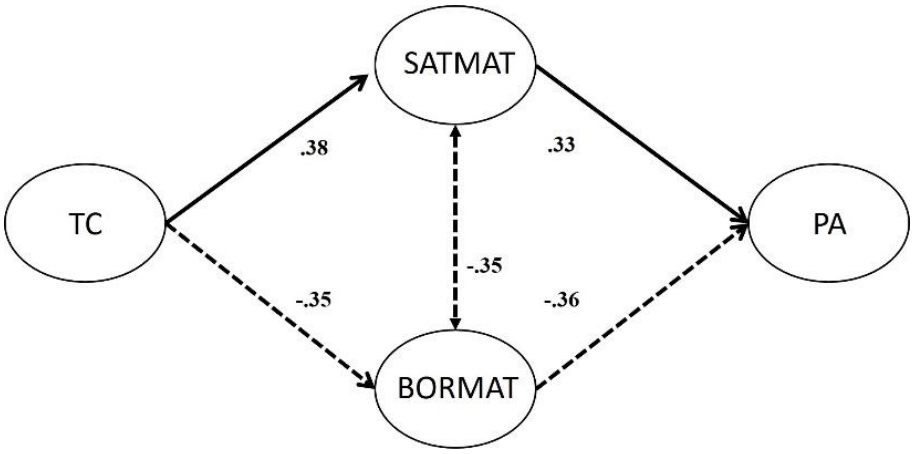
Structural regression modeling was used to analyze the existing relationships and interactions between the referred to theories. Several authors (Levy & Hankock, 2007) recommend formulating and analyzing various models (when data suggest doing so) and then reporting the most relevant results, therefore, various structural equation models were tested. Firstly, and following our initial hypothesis (see Figure 2), we estimated model A, in which the students' perception of teacher competencies predicts satisfaction with MAT, and in turn the scores in said subject. However, a perception of incompetence from the teachers will predict dissatisfaction with MAT, and this in turn a low performance in the subject.

Nonetheless, the obtained fitness indices yielded good results (χ^2 = 613.983; gl = 165; χ^2/gl = 3.72; p < 0.000; χ^2/gl = 3.72; NFI = 0.90; NNFI = 0.90; CFI = 0.92; RMSEA = 0.06, see Figure 2).

In view of such acceptable results, we decided to estimate a new model according to the indications suggested by the program, adding a direct relationship between the students' perception of teaching competencies and their academic performance in MAT. The results obtained from this new model B (see Figure 3) were excellent in several parameters: χ^2 = 524.837; gl = 164; χ^2/gl = 3.20; p < 0.000; χ^2/gl = 3.20; NFI = 0.92; NNFI = 0.93; CFI = 0.95; RMSEA = 0.05.

Figure 2

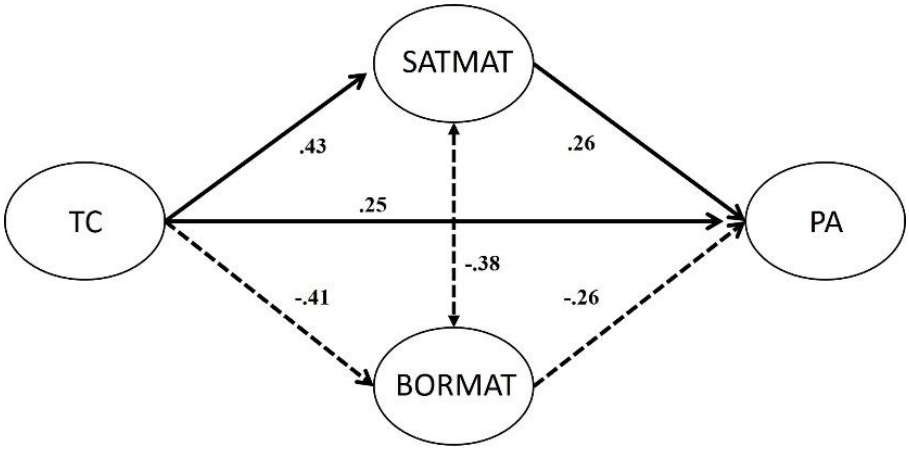
Structural Equation Model A



Note: TC = Teaching Competences; SATMAT = Satisfaction with Maths; BORMAT = Insatisfaction with Maths; PA = Performance academic with Maths.

Figura 3

Structural Equation Model B



Note: TC = Teaching Competences; SATMAT = Satisfaction with Maths; BORMAT = Insatisfaction with Maths; PA = Performance academic with Maths.

Discussion

According to the analysis of results, all four suggested hypotheses can be considered valid: the CCD instrument obtained adequate validity and reliability values in a sample with Mexican secondary school students (H1); factorial invariance had adequate values according to sex (H2); when students perceive their MAT teachers as competent, this predicts satisfaction with school and academic

performance (H3); and when students perceive their MAT teachers as incompetent, this predicts dissatisfaction with school and low academic performance (H4).

In terms of the first objective, the results of this unprecedented research proved the internal consistency and validity of the unidimensional CCD-MAT scale in a sample of Mexican secondary school students. Evidence of reliability and validity indices was provided, as well as data obtained from the goodness-of-fit analysis of the factorial structure through exploratory and confirmatory methods. In this regard, Baena-Extremera et al. (2015) obtained comparable results from their adaptation and validation of the instrument for physical education on Spanish secondary school students. It is worth underscoring that the CCD from Physical Education has been used in other research efforts as well, obtaining excellent results for instrument reliability and validity (Baños et al., 2019a; Baños et al., 2018; Granero-Gallegos et al., 2020), which demonstrates the scale's solidity in the Physical Education class. Therefore, the results obtained for the MAT class in this study indicate that the CCD-MAT scale is fully reliable and valid to measure the perceptions that students have of their MAT teachers' competencies, making a major contribution to scientific literature and to the assessment of the Mexican education system.

Regarding the second objective, this study revealed that the CCD-MAT is invariant according to sex. Similar results were seen (Granero-Gallegos et al., 2020) with Spanish secondary school students for the Physical Education class, underlining that the instrument works the same regardless of the subjects (EF and MAT). It is important to underscore that the CCD-MAT obtained a good fit in the factorial invariance analysis, provided that no studies are known to have analyzed invariance according to sex in an instrument assessing MAT teachers' competencies. The few studies that have found differences according to sex have analyzed variables such as support provided by the teacher (Riegle-Crumb, Farkas, & Muller, 2006), academic performance in MAT (Gherasim et al., 2013) and sentiment toward the subject (Frenzel et al., 2007); therefore, this study provides an invariant instrument to measure teaching competencies in MAT, deepening in the knowledge toward new lines of research.

Concerning the last objective, two prediction models were tested in this research: model A yielded an acceptable fit, whereas model B had an excellent fit. The prediction model presented in this study revealed, on the one hand, that when students perceive their MAT teachers as competent, MAT scores are positively predicted. Similar results were found by other researchers also regarding the MAT subject (Gherasim et al., 2013; Samuelsson & Samuelsson, 2016). A possible explanation is that when adolescents perceive that their teachers convey their knowledge comprehensibly, they feel heard during class, their interaction with the teacher increases, and their participation improves, or at least they show greater involvement in the classroom activities. Moreover, students perceiving their teacher as skilled positively predicts satisfaction with MAT, and in turn academic performance in MAT. On the contrary, when the student thinks that their MAT teacher is incompetent, dissatisfaction with MAT and low scores are predicted. In this regard, several studies carried out in Spain and Mexico have underscored the importance of the students feeling satisfied with the MAT subject, as this predicts satisfaction with school and, in turn, academic performance; accordingly, dissatisfaction with the subject predicts boredom at school and low academic performance (as demonstrated by Baños et al., 2019c; Baños et al., 2020). However, a concerning aspect seen in these studies is that Mexican students experienced higher levels of dissatisfaction with MAT and with school, which might seriously undermine their learning process and academic performance. Henceforth, dissatisfaction with Math does not only impact satisfaction with school directly but might also have an indirect effect on academic performance, provided that several studies have found that satisfaction with school predicts academic performance (Martin et al., 2015; Navarro et al., 2014). Nonetheless, boredom at school is associated with lower scores (Baños et al., 2020; Jyoty & Devi, 2008). In this manner, results obtained in this research provide relevant information for the Mexican education system in view of such low scores obtained in the PISA Report and PLANEA, demonstrating that teachers' competencies such as communication, work awareness, creativity, feedback, individual consideration, professionalism, problem solving and social awareness (all measured by the CCD-MAT) must be reworked to improve the adolescents' performance in MAT.

Conclusion

As a conclusion, the findings of this study proved that the CCD-MAT is a valid and reliable instrument to apply to Mexican secondary school students regardless of their sex. It is also worth underscoring those students perceiving their MAT teacher as competent predicts satisfaction with MAT and academic performance. On the contrary, a perception of incompetence predicts dissatisfaction with MAT and poor grades in the subject.

It is worth underscoring the strengths of this study, among which we emphasize the statistical analyses and the sample design, which was probabilistic and randomized in the centers, stratified, multistage and by proportional affixation. By doing so, the results of the study can be generalized for Mexico. Moreover, the topic addressed may help to remediate the low performance seen in Mexican adolescents during international evaluations. However, it also comes with a series of limitations that must be mentioned. The CCD-MAT may be used in secondary schools regardless of sex but shall not be used at other educational levels such as elementary school or high school.

Practical Implications

The CCD-MAT validation with secondary school students can have a positive impact on the learning process of MAT in the Mexican education system by being capable of analyzing the skills that MAT teachers need to improve, since it is their own students who evaluate them using a valid and reliable instrument. In this way, teachers will identify the skills in need of improvement that will subsequently have an impact on the student's satisfaction and their MAT grades, and that may even improve the scores obtained in future PISA Reports at the international level and in PLANEA at the national level.

Funding: This work had the support of the Secretary of Public Education of Mexico (identification number: 431/569/E) and belongs to the research project called "Program for International Student Assessment: relationship between school performance in secondary school students and the psychological, family and variable impact of physical activity".

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Ahmed, W., Minnaert, A., van der Werf, G., & Kuyper, H. (2010). Perceived social support and early adolescents' achievement: The mediational roles of motivational beliefs and emotions. *Journal of Youth and Adolescence, 39*(1), 36-46. https://doi.org/10.1007/s10964-008-9367-7

Ali, A., Tariq, R. H., & Topping, J. (2009). Students' perception of university teaching behaviours. *Teaching in Higher Education, 14*(6), 631-647. https://doi.org/10.1080/13562510903315159

Anaya, G. (1999). College impact on student learning: Comparing the use of self-reported gains, standardized test scores, and college grades. *Research in Higher Education, 40*(5), 499-526. https://doi.org/10.1023/A:1018744326915

Arbuckle, J.L., 2015. *Amos (Version 24.0)* [Computer Program]. IBM SPSS.

Baena-Extremera, A., Granero-Gallegos, A., & Martínez-Molina, M. (2015). Validación española de la Escala de Evaluación de la Competencia Docente en Educación Física de secundaria. *Cuadernos de Psicología del Deporte, 15*(3), 113-122. https://doi.org/10.4321/S1578-84232015000300011

Baños, R. Baena-Extremera, A., Ortiz-Camacho, M. M., Zamarripa, J., Beltrán, A., & Juvera- Portilla, J. L. (2019). Influencia de las competencias del profesorado de secundaria en los comportamientos disruptivos en el aula. *Espiral. Cuadernos del Profesorado, 12*(24), 3-10. https://doi.org/10.25115/ecp.v12i24.2141

Baños, R., Baena-Extremera, A., & Granero-Gallegos, A. (2019c). The Relationships between High School Subjects in terms of School Satisfaction and Academic Performance in Mexican Adolescents. *International Journal Environmental Research Public Health, 16*(18), 3494-3507. https://doi.org/10.3390/ijerph16183494

Baños, R., Baena-Extremera, A., & Ortiz-Camacho, M. D. M. (2019b). Prediction of the satisfaction with the student life, based on teaching competence and satisfaction with the school. *Frontiers in Psychology, 10*, 2506. https://doi.org/10.3389/fpsyg.2019.02506

Baños, R., Baena-Extremera, A., & Ortiz-Camacho, M. D. M. (2020). Prediction Model of Academic Performance and Satisfaction With School According to Some Subjects of Compulsory Secondary Education. *Psychological Reports, 123*(2), 435-451. https://doi.org/10.1177/0033294118805004

Baños, R., Barretos-Ruvalcaba, M., & Baena-Extremera, A. (2019a). Protocolo de estudio de las variables académicas, psicológicas y de actividad física que influyen en el rendimiento académico de adolescentes mexicanos y españoles. *Espiral. Cuadernos del Profesorado, 12*(25), 89–99. https://doi.org/10.25115/ecp.v12i25.2480

Baños, R., Ortiz-Camacho, M. M., Baena-Extremera, A., & Tristán-Rodríguez, J. L. (2017). Satisfacción, motivación y rendimiento académico en estudiantes de Secundaria y Bachillerato. *Espiral. Cuadernos del Profesorado, 10*(20), 40–50. https://doi.org/10.25115/ecp.v10i20.1011

Baños, R., Ortiz-Camacho, M. M., Baena-Extremera, A., & Zamarripa, J. (2018). Efecto del género del docente en la importancia de la Educación Física, clima motivacional, comportamientos disruptivos, la intención de práctica futura y rendimiento académico. *Retos, 33*, 252-257. https://doi.org/10.47197/retos.v0i33.59991

Baños, R., Toval, A., Morales-Delgado, N., & Ferrán, J. L. (2021). Analysis of movement during climbing as a strategy for learning the anatomy of the locomotor system in Sport Sciences. *Espiral. Cuadernos del Profesorado, 14*(29), 89-99. https://doi.org/10.25115/ecp.v12i25.2480

Barret, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences, 42*, 815-824. https://doi.org/10.1016/j.paid.2006.09.018

Bentler, P. M. (2007). *EQS for Windows* (v. 6.1). Multivariate Software.

Benton, S. L., Duchon, D., & Pallett, W. H. (2013). Validity of student self-reported ratings of learning. *Assessment & Evaluation in Higher Education, 38*(4), 377-388. https://doi.org/10.1080/02602938.2011.636799

Bollen, K. A. & Long, J. (1994). *Testing structural equation models*. Sage.

Carretero-Dios, H. & Perez, 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.

Catano, V. M. & Harvey, S. (2011). Student perception of teaching effectiveness: development and validation of the Evaluation of Teaching Competencies Scale (ETCS). *Assessment and Evaluation in Higher Education, 36*(6), 701-717. https://doi.org/10.1080/02602938.2010.484879

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

Douglass, J. A., Thomson, G., & Zhao, C. M. (2012). The learning outcomes race: The value of self-reported gains in large research universities. *Higher education, 64*(3), 317-335. https://doi.org/10.1007/s10734-011-9496-x

Espeland, V., & Indrehus, O. (2003). Evaluation of students' satisfaction with nursing education in Norway. *Journal of Advanced Nursing, 42*(3), 226-236. https://doi.org/10.1046/j.1365-2648.2003.02611.x

Fawad, H., & Manarvi, I. A. (2014, December). Student feedback & systematic evaluation of teaching and its correlation to learning theories, Pedagogy & Teaching skills. In 2014 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE) (pp. 398-404). IEEE. https://doi.org/10.1109/TALE.2014.7062572

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics—A "hopeless" issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education, 22*(4), 497-514. https://doi.org/10.1007/BF03173468

Gherasim, L. R., Butnaru, S., & Mairean, C. (2013). Classroom environment, achievement goals and maths performance: Gender differences. *Educational Studies, 39*(1), 1-12. https://doi.org/10.1080/03055698.2012.663480

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), 1-9. https://doi.org/10.3389/fpsyg.2020.00063

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate Data Analysi*s (7th ed.). Pearson Prentice Hall.

Helsinki Declaration (2008). *World Medical Association*. http://www.wma.net/es/30publications/10policies/b3/.

Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodology of investigation* (6thª ed.). México: McGraw-Hill.

Hodgson, Y., Varsavsky, C., & Matthews, K. E. (2014). Assessment and teaching of science skills: whole of programme perceptions of graduating students. *Assessment & Evaluation in Higher Education, 39*(5), 515-530. https://doi.org/10.1080/02602938.2013.842539

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.

Hu, L. & Bentler, P. M. (1999). *Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria*, https://doi.org/10.1080/10705519909540118

Jyoti, P., & Devi, P. N. (2008). Achievement motivation and its impact on academic stress study habits and academic performance among high school students. *International Social Science Journal, 24*(2), 107–115. https://doi.org/10.1037/t24331-000

Levy, R., & Hancock, G. R. (2007). A framework of statistical tests for comparing mean and covariance structure models. *Multivariate Behavioral Research*, *42*(1), 33-66. https://doi.org/10.1080/00273170701329112

Lidice, A. & Saglam, G. (2012). Using student evaluations to measure educational quality. *Procedia-Social and Behavioral Sciences, 70*, 1009-1015. 24. https://doi.org/10.1016/j.sbspro.2013.01.152

Lizzio, A., Wilson, K., & Simons, R. (2002). University students' perceptions of the learning environment and academic outcomes: implications for theory and practice. *Studies in Higher education, 27*(1), 27-52. https://doi.org/10.1080/03075070120099359

Luna, E., Calderón, N., Caso, J. & Cordero, G. (2012). Desarrollo y validación de un cuestionario de evaluación de la competencia docente con base en la opinión de los estudiantes. In E. J. Cisneros-Cohernour, B. García-Cabrero, E. Luna, and R. Marín (coords.), *Evaluación de Competencias Docentes en la Educación Superior* (pp.119-158). Juan Pablos Editor. 25.

Lynn, M. (1986). Determination and quantification of content validity. *Nursing Research*, *35*(6), 382-385. https://doi.org/10.1097/00006199-198611000-00017

Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry and J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319–384). Springer, https://doi.org/10.1007/1-4020-5742-3_9

Marsh, H. W., Martin, A. J., & Cheng, J. H. (2008). A multilevel perspective on gender in classroom motivation and climate: Potential benefits of male teachers for boys?. *Journal of Educational Psychology*, *100*(1), 78-95. https://doi.org/10.1037/0022-0663.100.1.78

Martin, A. J., Yu, K., Papworth, B., Ginns, P., & Collie, R. J. (2015). Motivation and engagement in the United States, Canada, United Kingdom, Australia, and China: Testing a multi-dimensional framework. *Journal of Psychoeducational Assessment*, *32*(2), 103–114. https://doi.org/10.1177/0734282914546287

Martínez-Molina, M., Granero-Gallegos, A., Baena-Extremera, A., & Baños, R. (2020). Invarianza factorial por sexo del cuestionario para las conductas disruptivas y escala de evaluación de la competencia docente en educación física en estudiantes de secundaria. *Psychology, Society, & Education*, *12*(2), 125-137. https://doi.org/10.25115/psye.v12i2.3172

McNaught, C., Ng, S. S., & Chow, H. (2012). Literacies in the humanities: the student voice. *Higher Education Research & Development, 31*(2), 139-154. https://doi.org/10.1080/07294360.2010.542558

Milfont, T. L. & Fisher, R. (2010). Testing measurement invariance across groups: Applications in cross. *International Journal of Psychological Research, 3*(1), 111–121. https://doi.org/10.21500/20112084.857

Muñiz, J., Elosua, P., & Hambleton, R.K. (2013). Directrices para la traducción y adaptación de los tests: Segunda edición. *Psicothema*, *25*(2), 151–7. https://bit.ly/3fmiJU6

Navarro, E., Expósito, E., López, E., & Thoilliez, B. (2014). EPIBI: Perception Scale of ChildWell-being Indicators. Validation of the instrument using polytomous Rasch models. *Revista de Educación, 364*, 39–65. https://doi.org/10.4438/1988-592X-RE-2014-364-254

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

OECD (2013). *Teaching and Learning International Survey TALIS 2013. Conceptual Framework.* Retrevied from: http://www.oecd.org/education/school/TALIS%20Conceptual%20Framework_FINAL.pdf

OECD (2019). *PISA 2018 Results. Where All Students Can Succeed*. Volume II. Paris: OECD Publishing. December 29, 2020,https://doi.org/10.1787/b5fd1b8f-en

Paisey, A., Kobayashi, H., & Li, J. (2007). Behavioural strategies of teachers in Japan. *Research in Education, 77*(1), 77-91. https://doi.org/10.7227/RIE.77.6

Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology, 99*(1), 83-98. https://doi.org/10.1037/0022-0663.99.1.83

Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology, 98*(3), 583-597. https://doi.org/10.1037/0022-0663.98.3.583

PLANEA. (2017). *Plan Nacional para la Evaluación de los Aprendizajes. Resultados nacionales 2017 Educación Media Superior*. INEE.

http://planea.sep.gob.mx/content/general/docs/2017/ResultadosNacionalesPlaneaMS2017.PDF

Puklek-Levpušček, M., & Zupančič, M. (2009). Math achievement in early adolescence: The role of parental involvement, teachers' behavior, and students' motivational beliefs about math. *The Journal of Early Adolescence, 29*(4), 541-570. https://doi.org/10.1177/0272431608324189

Rasmussen, J. F., Scrabis-Fletcher, K., & Silverman, S. (2014). Relationships among tasks, time, and student practice in elementary physical education. *Physical. Education, 71*, 114–131.

Riegle-Crumb, C., Farkas, G., & Muller, C. (2006). The role of gender and friendship in advanced course taking. *Sociology of Education, 79*(3), 206-228. https://doi.org/10.1177/003804070607900302

Samuelsson, M., & Samuelsson, J. (2016). Gender differences in boys' and girls' perception of teaching and learning mathematics. *Open Review of Educational Research, 3*(1), 18-34. https://doi.org/10.1080/23265507.2015.1127770

Shim, S. S., Ryan, A. M., & Anderson, C. J. (2008). Achievement goals and achievement during early adolescence: Examining time-varying predictor and outcome variables in growth-curve analysis. *Journal of Educational Psychology, 100*(3), 655-671. https://doi.org/10.1037/0022-0663.100.3.655

Shomoossi, N. (2004). The effect of teachers questioning behaviour on EFL Classroom interaction: A classroom research study. *The Reading Matrix, 4*(2), 96-104.

Simpson, P. M. & Siguaw, J. A. (2000). Student evaluations of teaching: An exploratory study of the faculty response. *Journal of Marketing Education, 22*, 199-214. https://doi.org/10.1177/0273475300223004

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

Tabachnick, B. G. & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). New York: Allyn and Bacon.

Takakura, M., Wake, N., & Kobayashi, M. (2010). The contextual effect of school satisfaction on health-risk behaviors in Japanese high school students. *Journal of School Health 80*, 544–551. https://doi.org/10.1111/j.1746-1561.2010.00540.x

Weber, K., Martin, M., & Patterson, B. (2001). Teacher behavior, student interest and affective learning: Putting theory to practice. *Journal of Applied Communication Research, 29*(1), 71-90. https://doi.org/10.1080/00909880128101