

Validación en español del Questionnaire on Teacher Interaction en los cursos superiores de educación primaria (*QTI-P*) y cómo esta interacción influye en el rendimiento académico

Francisco J. García Bacete^{1,2,*}, y Jesús F. Rosel Remírez¹

¹ Universitat Jaume I, Castellón (España).

² Miembro del Grupo GREI (Grupo interuniversitario de investigación del rechazo entre iguales en contextos escolares)

Resumen: The Questionnaire on Teacher Interaction (*QTI*, Wubbels et al., 1985) se basa en el Modelo Interpersonal de la Conducta del Profesor (*MITB*, en inglés) y mide las interacciones del maestro con el alumnado percibidas por éstos. Goh y Fraser (1998) adaptaron el *QTI* para su uso en los cursos superiores de la educación primaria (*QTI-P*). Los objetivos del estudio son (a) validar al castellano el *QTI-P* y (b) analizar la influencia de estas interacciones en el rendimiento académico del alumno. Se administró el *QTI-P* a 397 alumnos de 4º y 6º curso de cuatro colegios públicos seleccionados de forma incidental. Las 8 escalas del *QTI-P* se ordenan de forma circular alrededor de dos ejes o dimensiones independientes: Control y Afiliación. Como medida de rendimiento académico se usaron las calificaciones de final de curso en cuatro asignaturas. Para comprobar los objetivos se emplearon modelos de ecuaciones estructurales multinivel. La fiabilidad y la correlación intraclase de escalas y dimensiones son buenas. Se confirma la estructura factorial, circular y multinivel propuesta por el *MITB*. También se comprueba el efecto del ambiente social del aula tanto en la puntuación de cada niño en cada escala del *QTI-P* como en sus calificaciones escolares. El *QTI-P* es un instrumento adecuado para medir el ambiente social del aula y para pronosticar el rendimiento académico de cada niño.

Palabras clave: Relaciones interpersonales maestro-alumnado. Ambiente social del aula. Cursos superiores de educación primaria. Rendimiento académico.

Title: Spanish validation of the Questionnaire on Teacher Interaction in the upper grades of primary education (*QTI-P*) and how this interaction influences academic performance.

Abstract: The Questionnaire on Teacher Interaction (*QTI*, Wubbels et al., 1985) is based on the Model for Interpersonal Teacher Behaviour (*MITB*) and measures teacher-student interactions as perceived by students. Goh and Fraser (1998) adapted the *QTI* for use in the upper grades of primary education (*QTI-P*). The objectives of this work are (a) to validate the Spanish version of the *QTI-P*, and (b) to analyse the influence of these interactions on students' academic performance. The *QTI-P* was administered to 397 fourth and sixth grade students at four incidentally selected public schools. The eight scales of the *QTI-P* are arranged in a circular structure around two independent axes or dimensions: Control and Affiliation. End-of-year grades in four subjects were used as a measure of academic performance. Multilevel structural equation models were used to assess the objectives. We found suitable reliability and intraclass correlation of scales and dimensions. The factorial, circular and multilevel structure proposed by the *MITB* was validated. The effect of the classroom social environment on each child's score in each *QTI-P* scale and on their school grades was also verified. The *QTI-P* is a valid instrument to measure the social environment of the classroom and to predict the academic performance of each child.

Keywords: Teacher-student interpersonal relationships. Classroom social environment. Upper primary education. Academic performance.

Introduction

El concepto de ambiente de aprendizaje, desarrollado a partir de los trabajos pioneros de Murray y Lewin durante los años 30, reconoció que la conducta es función de la persona y del entorno. En el contexto escolar, este principio implica que las conductas de la maestra y las del alumnado¹ se influyen mutuamente, y que su interacción es muy importante para el aprendizaje del alumno (Abelló et al., 2020). En los últimos 50 años, este reconocimiento ha favorecido la investigación del ambiente de aprendizaje en el aula a través sobre todo de las percepciones de los estudiantes (Fraser y Walberg, 2005).

La investigación muestra que las variables docentes son

los mejores predictores de los resultados escolares de los estudiantes (den Brok y van Tartwijk, 2015). Se argumenta que la "interacción maestro-alumnado (*IMA*)" es uno de los principales factores que determinan su éxito escolar (Jimerson y Haddock, 2015; Roorda et al., 2011). Sin embargo, cada maestro muestra un comportamiento particular, que es diferente al de sus colegas y que los estudiantes pueden o no apreciar. Y también es cierto, que hay profesores a los que les gusta, o no, el comportamiento de sus estudiantes. Pero, más allá de gustos particulares, los estudiantes de una clase suelen ver el ambiente de aprendizaje de su aula de modo similar y diferente al de los estudiantes de otras clases (Scott y Fisher, 2004). Numerosos estudios han encontrado logros académicos superiores y actitudes más positivas de los estudiantes en las clases en las que éstos percibían comportamientos interpersonales más positivos en sus maestras (Cornelius-White, 2007; Tennant et al., 2015). Por el contrario, la percepción de relaciones conflictivas con los maestros, pueden conducir a una motivación más pobre y a un peor rendimiento (Wubbels et al., 2012). La teoría interpersonal ha conceptualizado las *IMA* fructíferamente (Wubbels et al., 2012), en términos de conductas de Afiliación o afectivas y de Control u organizativas de las maestras, concluyendo que cuanto más positivas son estas conductas, mejores son los resultados de aprendizaje de los estudiantes (Passini et al.,

* Correspondence address [Dirección para correspondencia]:

Francisco Juan García Bacete, Universitat Jaume I, Departamento de Psicología Evolutiva, Educativa, Social y Metodología, Avda Vicente Sos Baynat s/n, 1, 12071-Castellón (España). E-mail: fgarcia@uji.es
(Artículo recibido: 18-02-2020; revisado: 20-08-2020; aceptado: 28-08-2020)

¹ En España se usa de forma generalizada la denominación 'maestro' para referirse a los docentes que enseñan en educación infantil y primaria. A lo largo del texto utilizamos indistintamente ambos géneros, sin que ello signifique que nos estamos refiriendo a figuras distintas. No se renuncia tampoco a usar las expresiones más genéricas de profesor y profesora. Alternativamente, también recurrimos a la utilización de nombres colectivos como profesorado, docente o alumnado.

2015), de ahí la importancia de desarrollar instrumentos para conocer cómo los estudiantes perciben las conductas de sus maestras.

Nuestra meta es validar el *Questionnaire on Teacher Interaction (QTI)* para su uso con estudiantes de cursos superiores de educación primaria española. El *QTI* fue desarrollado originalmente por Wubbels, Créton y Hooymayers (1985) para examinar las *IMA* a través de las percepciones de los estudiantes de secundaria. Ha sido traducido y revisado en más de 15 lenguas y numerosos países (Fraser y Walberg, 2005). El *QTI* se ha aplicado muy poco en educación primaria (*QTI-P*; Goh y Fraser, 1998), sin embargo, la cultura de la escuela primaria es diferente de la de secundaria (Maulana et al., 2012), y los maestros de primaria crean ambientes de aula diferentes de los del profesorado de secundaria (Fisher et al., 2011), por lo que los resultados obtenidos con el *QTI* en educación secundaria no son directamente generalizables a primaria. Además, tampoco se pueden transferir los resultados de unos países a otros. Los estudios multiculturales y cross-nacionales con el *QTI* en secundaria (den Brok y van Tartwijk, 2015) muestran que, aunque los diferentes países suelen compartir la estructura del *QTI*, existe variabilidad en la percepción de la conducta del profesor de unos países a otros (Maulana et al., 2012). En consecuencia, es necesario avanzar en el conocimiento de las *IMA* del profesorado de educación primaria en diferentes países.

El modelo interpersonal de la conducta del profesor

Wubbels et al. (1985) aplicaron el modelo sobre las relaciones interpersonales propuesto por Leary (1957) al contexto de la educación. El modelo de Leary se representa como un círculo con dos dimensiones independientes que se cruzan: Control (Dominancia-Sumisión), el grado en que alguien domina el proceso de comunicación, y Afiliación (Cooperación-Oposición), el grado de cooperación que sienten los involucrados en el proceso de comunicación. Se pueden definir puntos intermedios entre Dominancia, Cooperación, Sumisión y Oposición para dividir el círculo interpersonal en 8 sectores.

El Modelo de la Conducta Interpersonal del Profesor (*MITB*, en inglés) propuesto por Wubbels et al. (1985, 2012) usa los mismos ejes de Control y Afiliación (dimensiones) que el de Leary (1957) y describe ocho tipos de conductas interpersonales (escalas) mostrados por los maestros (*IMA*): Directiva (E1), Amable (E2), Comprensiva (E3), Acomodatía (E4), Insegura (E5), Insatisfecha (E6), Represiva (E7), e Impositiva (E8) (Figura 1). El *MITB* se caracteriza por su estructura circunfleja o circular, que se basa en los siguientes supuestos: (1) las ocho escalas de conducta del *MITB* están representadas por dos dimensiones, Control y Afiliación; (2) las dos dimensiones no están correlacionadas; (3) las escalas se distribuyen uniformemente en una estructura circular. Se espera que cada escala cargue simultáneamente en ambas dimensiones, aunque con diferente magnitud. Las puntuaciones de las dimensiones son una combinación lineal de las

puntuaciones de las 8 escalas del *QTI* conforme a las siguientes ecuaciones, en donde los coeficientes ($\pm .92$ o $\pm .38$) son las coordenadas geométricas teóricas de cada escala en el círculo interpersonal (Wubbels y Brekelmans, 2005) (Figura 1):

$$\text{Control} = .92 * E1 + .38 * E2 - .38 * E3 - .92 * E4 - .92 * E5 - .38 * E6 + .38 * E7 + .92 * E8$$

$$\text{Afiliación} = .38 * E1 + .92 * E2 + .92 * E3 + .38 * E4 - .38 * E5 - .92 * E6 - .92 * E7 - .38 * E8, \quad (1)$$

siendo cada escala: E1, Directiva..., y E8, Impositiva. Obsérvese: (a) que Control y Afiliación están medidas sin error, y (b) que aquellas escalas cuya distancia es de 3 (E1 con E4, E2 con E5, ...) su correlación es igual a cero, (Figura 1) o, usando términos factoriales, son ortogonales (Méndrila, 2017).



Figura 1. Estructura circunfleja del modelo de la conducta interpersonal del profesor (Wubbels et al., 2012).

Antecedentes de la investigación con *QTI* en educación primaria (*QTI-P*)

El *QTI* se desarrolló en los Países Bajos para analizar las conductas interpersonales de los docentes por medio de las percepciones del alumnado de secundaria, utilizando ocho escalas que se corresponden con los sectores del *MITB*. Wubbels y Levy (1991) redujeron la versión original de 77 a 64 ítems (8 escalas con 8 ítems) para usarla con muestras estadounidenses. Fisher, Henderson y Fraser (1995) validaron una nueva versión de 48 ítems para su uso en el contexto educativo australiano (cada escala tiene 6 ítems). Varios estudios han encontrado valores satisfactorios de fiabilidad y validez en el *QTI* (Wubbels et al., 2012).

Goh y Fraser (1998) adaptaron la versión australiana del *QTI*-secundaria para su uso en cursos superiores de educación primaria de Singapur (*QTI-P*), utilizando vocabulario y

estructuras de oraciones más simples que el *QTI*-secundaria, y un formato de respuesta más sencillo, cambiando de “cinco” a “tres” puntos y de “importante” a “frecuente”. El *QTI-P* fue validado originalmente con 1512 estudiantes de 5º curso de 39 aulas de 13 escuelas seleccionadas al azar.

El *QTI-P* ha sido validado con alumnado de edades comprendidas entre 9 y 14 años en múltiples países, Malasia (Scott y Fisher, 2004), Brunei (den Brok et al., 2005), Grecia (Kokkinos et al., 2009) y Turquía (Telli y den Brok, 2012). Los resultados con *QTI-P* replican los obtenidos en secundaria, pero con valores algo más bajos de fiabilidad (Kokkinos et al., 2009; Telli y den Brok, 2012), y con medias más altas

en las dos dimensiones (Control y Afiliación), en particular en Afiliación (Fisher et al., 2011).

En general, estos estudios utilizan grandes muestras, se suelen concentrar en el rango alto de edad (quinto curso o superiores), usan escalas de respuesta de tres puntos para referirse a la conducta del docente de una materia determinada, siguen procedimientos similares para estudiar las características psicométricas del *QTI-P* y suelen emplear el rendimiento académico como criterio para el estudio de la validez predictiva. En la Tabla 1 se presenta una comparativa de los estudios que han empleado el *QTI-P*.

Tabla 1. Comparativa estudios previos sobre la versión del *QTI* para alumnos de los cursos superiores de educación primaria

Investigaciones ^a	Goh y Fraser (1998)			Scott y Fisher (2004)			den Brok, Fisher, y Kokkinos, Charalampous, Telli y den Brok (2012)			Scott (2005)			y Davazoglou (2009)							
QTI de partida	Australiana			Australiana			Australiana			Australiana			Americana							
País	Singapur			Malasia			Brunei			Grecia ^b			Turquía							
Muestra	1512			3194			1305			273			626							
Curso	5º			4º-5º-6º			6º-7º-8º			5º-6º			Sin especificar							
Escala de respuesta	3 puntos			3 puntos			3 puntos			3 puntos			3 puntos							
Escalas <i>QTI-P</i>	α	α	η^2	Ítem	α	α	η^2	Ítem	α	α	ICC	Ítem	α	α	η^2	Ítem	α	α	η^2	Ítem
	Ind	Cl		Elim ^c	Ind	Cl		Elim ^c	Ind	Cl		Elim ^c	Ind	Cl		Elim ^c	Ind	Cl		Elim ^c
Directiva	.63	.90	.18	---	---	.81	.20	0	---	.82	.24	0	.73	---	.13	1-9-25	---	.83	.20	---
Amable	.78	.96	.38	---	---	.84	.20	0	---	.83	.19	-1	.76	---	.12	2-34-42	---	.90	.18	---
Comprensiva	.65	.94	.30	---	---	.81	.18	0	---	.82	.22	-2	.64	---	.14	3-19-35	---	.70	.20	---
Acomodaticia	.58	.73	.13	---	---	.75	.18	0	---	.76	.26	0	.59	---	.44	4	---	.76	.26	---
Insegura	.50	.83	.15	---	---	.84	.21	37	---	.79	.14	0	.36	---	.33	Todos	---	.90	.20	---
Insatisfecha	.76	.96	.33	---	---	.77	.18	0	---	.74	.16	-1	.65	---	.12	6,22	---	.86	.16	---
Represiva	.74	.93	.35	---	---	.86	.24	0	---	.85	.22	0	.63	---	.19	23	---	.92	.28	---
Impositiva	.58	.81	.15	---	---	.62	.14	24	---	.70	.17	-2	.55	---	.23	8,16,24	---	.87	.31	---
Puntuaciones	Agregadas de Aula			Agregadas de Aula			Agregadas de Aula			Puntuaciones del alumno			Agregadas de Aula							
Factorización	No. Usa escalas origi- nales del <i>QTI</i> para Se- cundaria			No. Usa escalas origina- les del <i>QTI</i> para Secundaria			No. Usa escalas origi- nales del <i>QTI</i> para Se- cundaria			AFE (20 ítems, 5 escalas)			No. Usa escalas origina- les del <i>QTI</i> para Secundaria							
Estructura Circular	Correlaciones interes- calas			Correlaciones interescalas			Correlaciones interes- calas			Correlaciones interescalas			Correlaciones interescalas							
Subyacen 2 Dimensiones (Control y Afiliación)	---			---			---			AFE multinivel con escalas. Solo significa- tiva Control			---			AFE con escalas. 3 Dim Solo interpretable Afilia- ción				
Dimensiones Independientes Multinivel	---			---			---			---			---			Correlación Dimensiones				
Validez Predictiva	NO			NO			NO			SÍ			NO			NO				
	Criterio: Test rendi- miento y Escala califi- cación de la atracción en Matemáticas. Téc- nica: Correlación y MANOVA			Criterio: Test rendimiento y Escala calificación del disfrute en Ciencias: Téc- nica: Correlación y Regre- sión lineal			Criterio: Media en 7 ítems de disfrute en Ciencias Técnica: ANOVA multinivel (estudiante, aula, profesor), con Dimensiones			Criterio: Estimación del Rendimiento en Matemá- ticas y en Lengua juntas (3 niveles Técnica: ANOVA (Géne- ro, SES; Rendimiento)										

^aNo incluimos estudios que se han aplicado a estudiantes de primaria, pero que no han usado el *QTI-P*, sino la versión del *QTI* para secundaria (por ejemplo, Fisher et al., 2011, en Australia, y el de Poulu, 2014, en Grecia).

^bEn 2017, Charalampous y Kokkinos realizan un nuevo estudio en Grecia, pero tampoco lo hemos incluido porque incorporan 21 ítems no originales al *QTI-P*.

^cÍtems eliminados: No se ha realizado (---), o no se indica qué ítems concretos se han eliminado (-1, -2), o no se explica cómo se ha hecho.

Usaremos varias novedades respecto de la muestra, la escala de respuesta, el criterio, y los procedimientos de validación. Emplear muestras grandes puede sesgar algunos resultados, como el ajuste al modelo (Iacobucci, 2009). Para evitar que nuestros resultados estén afectados por características de la muestra, emplearemos una muestra más pequeña y, puesto que el objetivo es estudiar la respuesta de alumnos más jóve-

nes de los usados hasta ahora, incluiremos alumnado de cuarto curso, primer curso que se considera educación primaria superior. Dos novedades más se refieren a la escala de respuesta del *QTI-P* y de quién se informa. Como indicaron Telli y den Brok (2012), puede ser que la escala de 3 puntos no permita una diferenciación suficiente en la versión *QTI-P* de la escuela primaria, lo que puede producir correlaciones

inesperadas. Además, Zijlstra, Wubbels, Brekelmans y Koomen (2013) han mostrado que los estudiantes de primer y segundo curso tienen una buena comprensión de una escala de respuesta de 5 puntos en el *QTI-EP* (versión para los primeros cursos de educación primaria). García Bacete, Ferrá, Monjas y Marande (2014) también llegan a la misma conclusión usando una muestra española. Tradicionalmente, se pide a los alumnos que informen sobre la conducta de un maestro de una asignatura. Nosotros hemos seleccionado la figura del maestro tutor, que representa mejor la cultura de la escuela primaria que el profesorado de una asignatura, más típico de la cultura de secundaria (García Bacete, Ferrá et al., 2014; Maulana et al., 2012). El maestro-tutor es el que más tiempo pasa en el aula, quien imparte la mayor parte de las asignaturas y sobre quien recae la responsabilidad de establecer las normas del aula (Jimerson y Haddock, 2015). Respecto al criterio, habitualmente se emplean pruebas estandarizadas de rendimiento académico (Scott y Fisher, 2004), nosotros usaremos las calificaciones escolares, dado que las percepciones de los estudiantes de las *IMA* tienen mayor efecto sobre las notas que los tests estandarizados (Roorda et al., 2011).

Respecto de las características psicométricas, estos estudios informan que las escalas tienen valores de fiabilidad altos cuando se calculan usando las puntuaciones agregadas de aula. Sin embargo, en los estudios que usan las puntuaciones de los alumnos, esos coeficientes de fiabilidad disminuyen mucho. De hecho, tres escalas, Acomodaticia, Insegura e Impositiva suelen tener valores de α inferiores a .60, especialmente la escala Insegura. A pesar de estos resultados, no se suelen tomar decisiones sobre la composición de las escalas, quizás porque los valores de η^2 o los índices de correlación intraclase (*ICC*) son buenos, normalmente superiores a .15 en todas las escalas.

Respecto de la validación del modelo circunflejo, se observan dos tendencias generalizadas: (a) el uso de puntuaciones agregadas de aula en escalas y dimensiones, ignorando las puntuaciones individuales de los estudiantes, y (b) la utilización en los análisis factoriales de los efectos de las dimensiones sobre las escalas, ignorando los de las escalas sobre los ítems (den Brok, Brekelmans, y Wubbels, 2006), reduciendo los modelos factoriales de segundo orden a modelos de primer orden. Pero como señalan Lüdtke et al. (2008), las variables agregadas de aula solo pueden utilizarse a nivel 2, y los estudios que las han usado en un primer nivel de análisis han cometido errores de sesgo y de distribución de parámetros (F , t , z ,...). Los métodos más extendidos han sido el análisis del patrón de las correlaciones interescalas, que informan de un ajuste aceptable de las escalas a las dimensiones, con algún solapamiento entre escalas (por ejemplo, Telli y den Brok, 2012), y el análisis factorial exploratorio (*AFE*) (Kokkinos et al., 2009). No obstante, estos métodos presentan dificultades para confirmar las dos dimensiones, su ortogonalidad y la posición de las escalas (Telli y den Brok, 2012). Los únicos trabajos que han usado análisis factorial confirmatorio (*AFC*) han sido los realizados por Kokkinos et al. (2009), pero sólo confirma 7 escalas con α entre .51 y

.64, y por den Brok et al. (2005) que también era multinivel, pero usaron puntuaciones agregadas de aula y no puntuaciones de los alumnos. Nosotros realizaremos *AFC*, con factores de segundo orden, y multinivel, en los que se considerarán tanto las variables observables (ítems) como las escalas, las puntuaciones de los estudiantes (nivel 1) como las puntuaciones de aula (nivel 2).

No obstante, los estudios muestran la validez predictiva del *QTI-P* usando diferentes criterios (rendimiento académico, disfrute, en una o varias asignaturas). Los logros más altos y las mejores actitudes de los estudiantes en las aulas, el agrado y el disfrute en las asignaturas de matemáticas y ciencias, se encontraron en las aulas en las que las maestras enfatizaron las conductas Directivas, Amables y Comprensivas, y mostraron menos signos de conductas Insatisfechas y Represivas (Goh y Fraser, 1998; Kokkinos et al., 2009; Scott y Fisher, 2004). Se han usado diferentes procedimientos de análisis (correlaciones, regresión lineal, modelo lineal jerárquico, o análisis de varianza). Nosotros utilizaremos *AFC* y modelos de ecuaciones estructurales (*SEM*) multinivel.

Objetivos e hipótesis

Nuestro objetivo es comprobar cuál es la estructura (factorial, circunfleja y multinivel) del *QTI-P* en lengua castellana y, además, comprobar si el *QTI-P* es un buen predictor del rendimiento académico del alumnado, sobre la base de una minuciosa atención a la composición de las escalas y a las diferencias culturales, como sugieren Wubbels et al. (2012). Para ello, nuestro estudio plantea cuatro hipótesis ($H\#$): ($H1$) La estructura del *QTI-P* en castellano para niños de 4º a 6º curso de educación primaria sigue la estructura *factorial* del *QTI-P* en lengua inglesa. Esto es, las percepciones interpersonales de los alumnos de la interacción de la maestra con los alumnos, expresadas a través del *QTI-P*, se agrupan en 8 escalas conductuales (E1 a E8). ($H2$) Los datos del *QTI-P* en castellano se ajustan a la estructura *circunfleja* definida por el *MITB*, Ecuación 1. ($H3$) Como los alumnos están agrupados en aulas, los datos adoptarán una estructura *multinivel*, siendo el Nivel 1: alumno, y el Nivel 2: aula. Esto es, el ambiente del aula tendrá un efecto en las percepciones de los alumnos de las conductas del maestro. ($H4$) Las percepciones interpersonales de los alumnos sobre la conducta del maestro, dimensiones del *QTI-P*, *pronostican el aprendizaje escolar*, medido con las calificaciones obtenidas a final de curso por cada niño.

Método

Participantes

Este trabajo forma parte de un proyecto de investigación de mayor envergadura sobre el rechazo entre iguales en la educación primaria y las variables que intervienen en su desarrollo y mantenimiento y en la intervención para prevenir y corregirlo (García Bacete et al., 2019). El presente estu-

dio se realizó con todos los estudiantes de cuarto y sexto curso de educación primaria ($n = 409$) en cuatro escuelas públicas españolas, situadas en zona urbana, seleccionadas de forma incidental en dos localidades próximas a la Universidad que lideraba el proyecto, con alumnado de familias de nivel socioeconómico medio y de origen étnico mayoritariamente caucásico (97%). La participación fue del 97%, 397 estudiantes, 212 (edad media = 9.91; $DT = 0.28$; 55.7% niñas) escolarizados en 9 clases de cuarto curso (tamaño medio aulas = 23 alumnos-as) y 185 (edad media = 11.89; $DT = 0.38$; 45.4% niñas) en 10 clases de sexto curso (tamaño medio aulas = 19 alumnos-as). Los 12 alumnos que no participaron fue debido a que se acaban de incorporar a estas escuelas en los últimos meses, no tenían autorización parental o no asistieron a la escuela el día de administración del cuestionario. El rendimiento medio en las 4 materias evaluadas fue 3.70 y 3.46 en matemáticas, 3.78 y 3.50 en ciencias, 3.89 y 3.69 en lengua castellana y 3.70 y 3.58 en valenciano, en 4º y 6º respectivamente. La única diferencia significativa entre cuarto y sexto es en ciencias ($t = 2.21$, $gl = 395$, $p = .028$). Un 9.4% de los niños y niñas de cuarto tienen dificultades de aprendizaje en cuarto y un 6.4% en sexto curso. La distribución sociométrica es la misma en ambos cursos (78.3% y 73.7% de medios, 8.0% y 12.4% de preferidos, 9.9% y 9.7% de rechazados, 1.9% y 2.7% de ignorados, 1.9% y 1.1% de controvertidos, en 4º y 6º respectivamente) y se corresponde con los porcentajes medios encontrados por García Bacete, Jiménez et al. (2014) para la población española. De los 19 docentes tutores, los 9 de cuarto curso eran maestras, y de los 10 de sexto, 5 eran maestros y 5 maestras. 17 de los 19 venían actuando como tutores y tutoras en estas aulas desde principio de ciclo, y 2 maestras de 6º se incorporan al grupo al inicio de curso. Solo una maestra de cuarto curso era nueva en el colegio. Todos los tutores tienen una experiencia docente superior a 10 años, excepto una docente cuya experiencia era de 6 años.

Medidas

Questionnaire on Teacher Interaction-Primary (QTI-P). Los ítems del *QTI-P* (Goh y Fraser, 1998) son descripciones de conductas típicas del maestro con el estudiantado en el aula. Los ítems del *QTI-P* enfatizan percepciones compartidas, puesto que ponen al alumno en un nivel de aula, por ejemplo, preguntando “Este maestro es simpático”, mientras que este otro enunciado “Este maestro me parece simpático” pide la opinión individual del alumno. Los 48 ítems se distribuyen en ocho escalas con 6 ítems cada una. Los ítems en el cuestionario están organizados en orden cíclico, de modo que el primer, segundo, ... y octavo ítem en cada bloque evalúa una de las ocho escalas del modelo. En el Anexo se presentan los 48 ítems del *QTI-P* en castellano, indicando la escala a la que pertenecen (Goh y Fraser, 1998). Los ítems fueron adaptados del inglés al castellano mediante un proceso de traducción inversa con la participación de 5 expertos, uno de ellos nativo inglés. Se pidió a los niños que respondieran,

acerca de su maestro-tutor, en una escala Likert de 5-puntos, de 0 a 4 (“nunca ocurre”, “ocurre muy poco”, “ocurre a veces”, “ocurre muy a menudo” y “siempre ocurre”). La puntuación de cada alumno en cada escala del *QTI-P* es la media aritmética de sus puntuaciones en los ítems que forman la escala (rango 0-4); posteriormente, por razones de interpretación, se transformaron a una puntuación con rango 0-1, y valores continuos, lo cual no afecta a la estructura factorial, puesto que la correlación entre una escala 0-4 y otra 0-1 es igual a la unidad ($r = 1$), por lo que no cambia la estructura de covarianzas de los modelos, siendo idénticos los indicadores de ajuste y los de significación de los parámetros. Finalmente, la puntuación de cada alumno en cada una de las dos dimensiones se calcula aplicando las Ecuaciones 1, propuestas por Wubbels y Brekelmans (2005).

Rendimiento Académico. Se utilizaron las calificaciones escolares de matemáticas, ciencias, castellano y valenciano tomadas directamente de las actas de evaluación final de 4º y 6º curso. Se aplicó una escala de 5 puntos (insuficiente, suficiente, bien, notable y sobresaliente), con un rango de 1 a 5. Para calcular la puntuación a nivel 2 de rendimiento de cada asignatura, se ha hecho la media por clase a partir de cada respectiva asignatura para cada uno de los alumnos. El estudio de Kokkinos et al. (2009) también utilizó las calificaciones del rendimiento emitidas por el profesor, pero en su investigación usaron una estimación de las calificaciones y no las calificaciones reales, como se hace en nuestro estudio.

Procedimiento

Este estudio se realizó con aprobación de la Comisión Deontológica de la Universidad, los Consejos Escolares de las escuelas y la Consellería de Educación de la Generalitat Valenciana (España). El *QTI-P* se administró colectivamente en cada aula, durante un tiempo aproximado de 20 minutos. La participación fue voluntaria. Las familias dieron consentimiento informado escrito.

Estrategia Analítica

Los análisis de datos se llevaron a cabo con los programas SPSS 24 (IBM, 2016) y EQS 6.3 (Bentler, 2016). Con el fin de comprobar las hipótesis, se han realizado análisis descriptivos, de fiabilidad, correlacionales (Pearson e *ICC*), *AFC* y *SEM* multinivel. En los *AFC* y *SEM*, como criterios de ajuste se han utilizado los índices propuestos por Steiger (2007). Para la comparación entre modelos no anidados utilizamos el procedimiento de Akaike Information Criterion (*AIC*; Burnham, y Anderson, 2004).

Con el fin de comprobar la primera hipótesis (H1), sobre la composición de cada escala, se han hecho *AFC* de primer orden (Figura 2) y análisis de fiabilidad para comprobar la consistencia interna de las 8 escalas. Primeramente, se comprobó si los ítems del *QTI-P* saturaban de manera significativa en cada uno de sus correspondientes factores de primer

orden, o escalas. Para ello, se utilizaron las puntuaciones individuales de los alumnos en los 48 ítems originales y covarianzas entre las 8 escalas (M1), hasta alcanzar una configuración

factorial que ajustara a los datos (M2). Para cada una de las escalas se estimó su media, desviación típica y α de Cronbach.

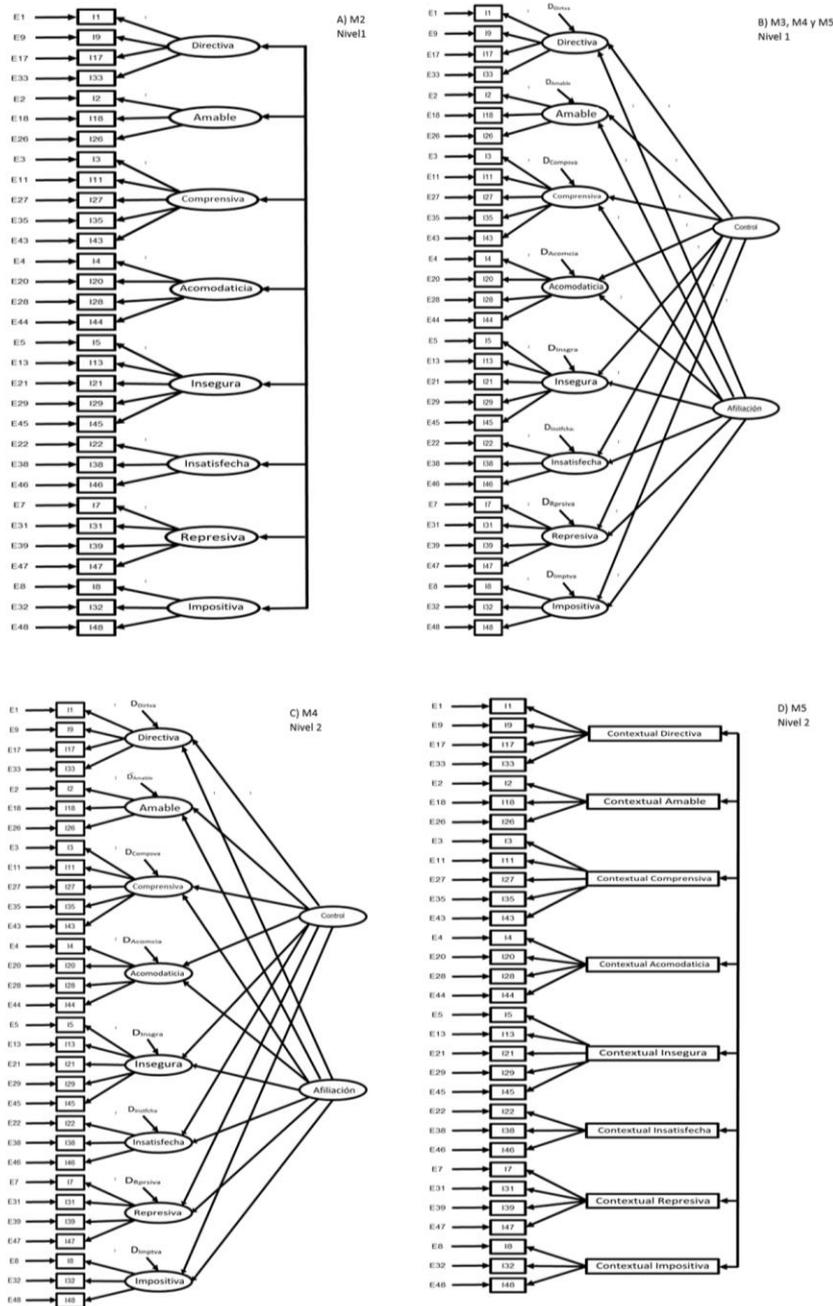


Figura 2. Modelos analizados. A) Modelo M2, sólo de nivel 1, los efectos fijos entre factores y variables son iguales a 1. B) Modelos M3, M4 y M5, sólo de nivel 1. Para los efectos fijos entre factores se han utilizado los coeficientes circunflejos. C) M4, nivel 2. Los efectos fijos son iguales a 1. D) es el nivel 2 del modelo M5.

Notas. M4 es el modelo del apartado B), para nivel 1, más el apartado C), para el nivel 2. M5 está integrado por los apartados B) y D).

En M3, M4 y M5 se han fijado los efectos de las dimensiones a las escalas mediante los coeficientes circunflejos, Ecuación (1).

f : efectos fijos, iguales a la unidad o a los efectos circunflejos, según cada caso.

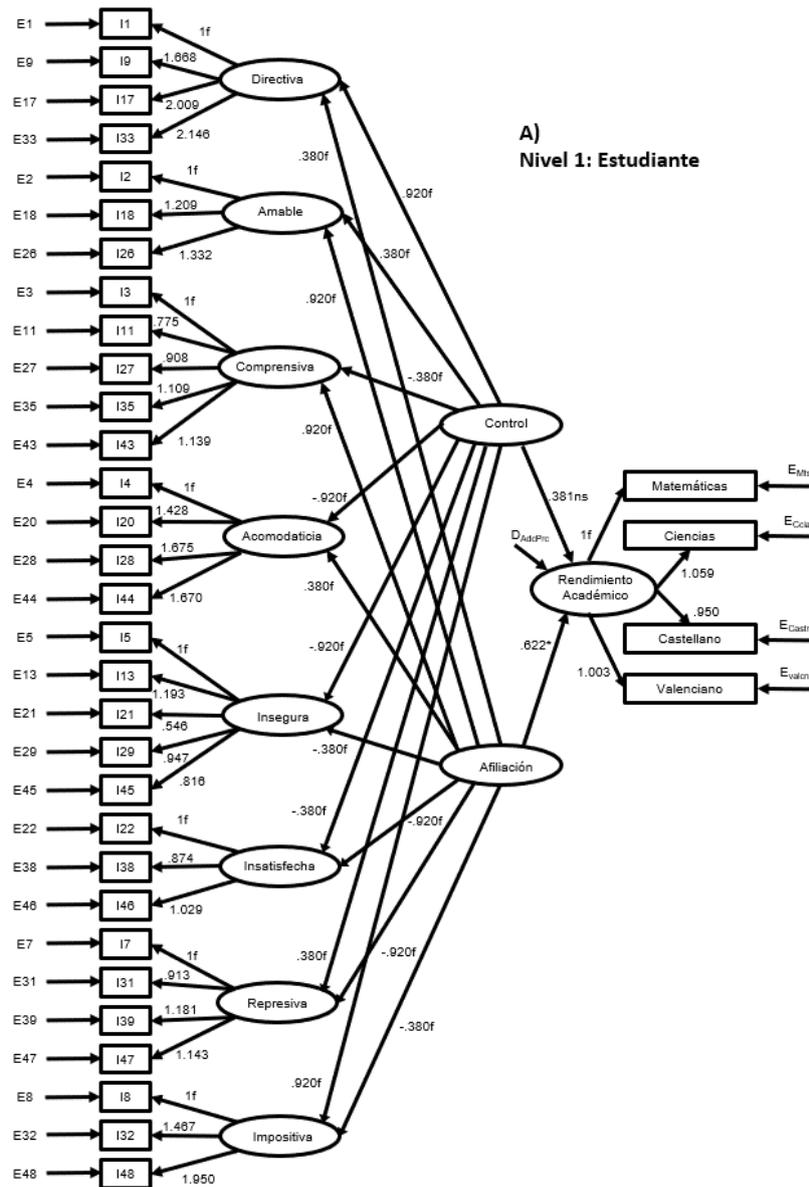
Para comprobar H2, sobre la estructura circunfleja del modelo, se usó M3. En la Figura 2, apartado B, hay una re-

presentación de M3. Este modelo comprende dos factores ortogonales de segundo orden, o dimensiones, que son las

variables independientes del modelo. Estas dimensiones, a través de las correspondientes cargas fijas factoriales circunflejas, influyen en las ocho escalas o factores de primer orden, que a su vez influyen en los ítems del *QTI-P*, de acuerdo con los supuestos del *MITB*. M3 incluye valores sólo a nivel 1, de alumno.

Al objeto de comprobar H3, sobre la influencia del aula (multinivel) en las percepciones de cada ítem del *QTI-P*, se calcularon los *ICC* de escalas y dimensiones, y se realizaron tres *AFC* multinivel. En los tres *AFC*, el primer nivel coincide con el de M3. A nivel 2, en el primer *AFC* se ha reproducido la estructura circunfleja de nivel 1 con efectos fijos, y en el segundo *AFC* se atenuó el modelo, dejando efectos li-

bres (M4), Figura 2, apartado B para el nivel 1 y apartado C para el nivel 2, conforme al sistema de representación de Muthén y Muthén (2017). Otra comprobación complementaria de esta hipótesis multinivel (M5), ha sido incluir variables contextuales en el nivel 2 (es decir, no se reproduce la estructura circunfleja), sino las respectivas medias de los valores de *QTI-P* para cada escala en cada aula (Fisher et al., 2011; Hox, Moerbeek, y Schoot, 2017; Lüdtke et al., 2008). Puesto que estas variables contextuales son valores proporcionados por los resultados (observables), no es preciso fijar a 1 ningún coeficiente de nivel 2, ver Figura 2, apartados B y D.



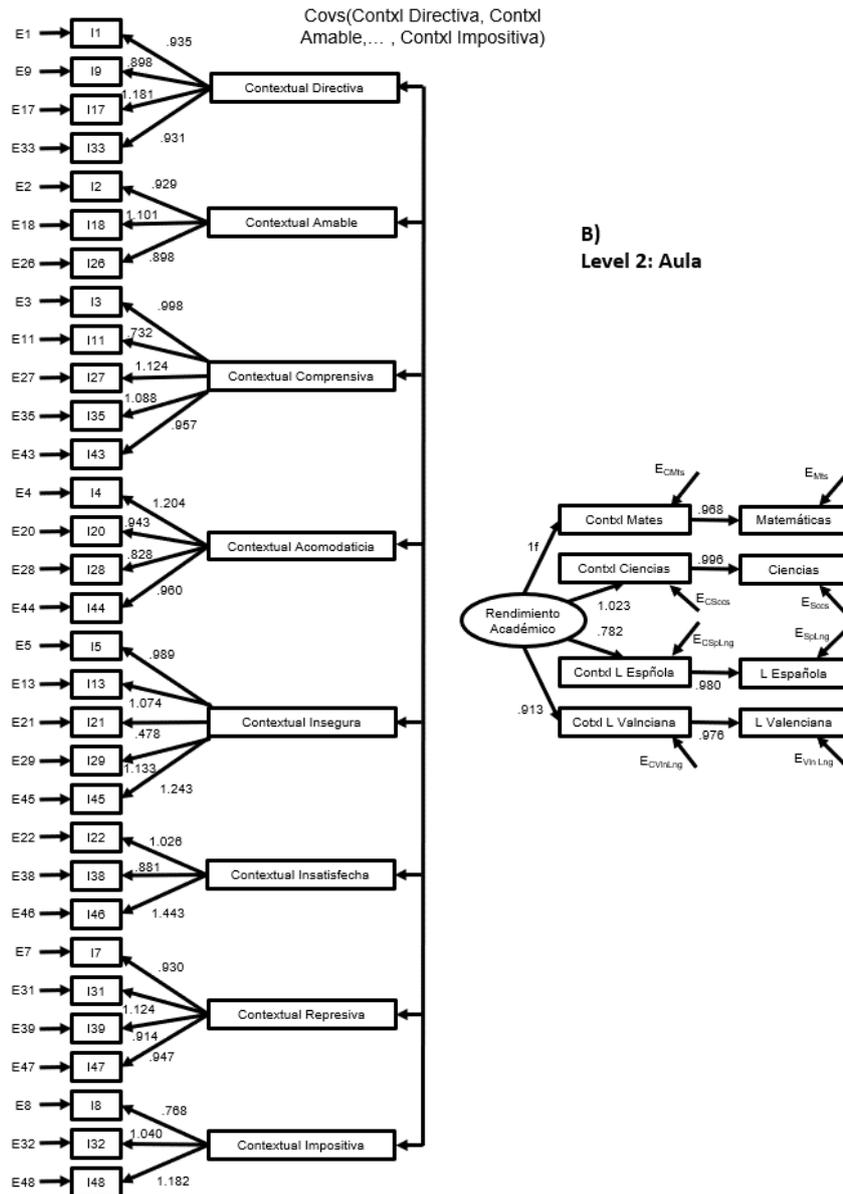


Figura 3. Estructura circunfleja del QTI-P confirmada en puntuaciones directas, incluyendo los efectos sobre el rendimiento académico de cada alumno (M6): A) para el nivel 1, alumno; y B) para nivel 2, Aula.

Nota. Para simplificar la figura, no se han incluido los valores de todas las posibles covarianzas entre los factores: Directiva, Amable, ..., e Impositiva, aunque están incluidas en el modelo, todas con $p < .001$.

* $p < .05$. f , efectos fijos. n , efecto no significativo. Todos los demás efectos tienen una $p < .001$.

Con el fin de comprobar H4, referida a la capacidad predictiva de las percepciones de los estudiantes medidas con el QTI-P sobre el rendimiento académico, se usaron como criterio de rendimiento académico a nivel 1 las calificaciones escolares del alumno en Matemáticas, Ciencias, Castellano y Valenciano a final de curso, y a nivel 2 la media de la respectiva calificación por aula, M6, y Figura 3. M6 se basa en M5, pero añadiendo los efectos sobre el rendimiento académico como acabamos de explicar. M6 es un modelo SEM completo.

Resultados

Las medias y desviaciones típicas de las escalas y de las dimensiones se presentan en la Tabla 2. La fiabilidad compuesta de las escalas oscila entre .60 en la escala Insegura y .88 en la escala Represiva. El rango del porcentaje de varianza medido con la varianza media extractada (VME) va entre 23.7% y 58.2%, para las escalas Insegura y Amable, respectivamente. En el caso de las dimensiones, Afiliación explica un 64.1% y Control un 2.27%.

Tabla 2. Información descriptiva (n=397) de las escalas y dimensiones interpersonales y del rendimiento académico (nivel alumno).

	<i>M</i>	<i>DT</i>	α	<i>FC</i> ¹	<i>VME</i> ¹	<i>ICC</i>
Directiva	.78	.16	.68	.607	.2918	.330
Amable	.79	.23	.86	.806	.5817	.357
Comprensiva	.76	.19	.83	.798	.4440	.243
Acomodatícia	.45	.19	.70	.605	.2808	.263
Insegura	.13	.15	.63	.602	.2374	.095
Insatisfecha	.18	.21	.75	.671	.4107	.404
Represiva	.27	.24	.87	.875	.4810	.466
Impositiva	.36	.25	.73	.617	.3630	.328
Control ²	.56	.34	--	--	.0227	.352
Afiliación ²	1.28	.88	--	--	.6405	.422
Matemáticas	3.59	1.25				.090
Ciencias	3.65	1.27				.112
Castellano	3.80	1.12				.059
Valenciano	3.64	1.19				.118

FC: Fiabilidad Compuesta; *ICC*: Coeficiente de Correlación Intraclass; *VME*: Varianza Media Extractada.

¹ Se han usado los pesos estandarizados de M5; el resto de los parámetros son válidos para todos los Modelos. ²Puntuaciones teóricas

(H1) El primer *AFC* (M1) se realizó con los 48 ítems originales del *QTI-P* y covarianzas entre todas las escalas. M1

Tabla 3. Modelos analizados y estadísticos de ajuste de conjunto.

Modelo	Descripción	χ^2	gl	<i>p</i> (χ^2)	χ^2 /gl	NNFI	CFI	IFI	SRMR	RMSEA	AIC
M1	<i>QTI-P</i> 45 ítems, Nivel 1, 8 factores 1 ^{er} orden, con Covarianzas	2233.9	1052	<.001	2.12	.840	.851	.853	.062	.056	129.9
M2	<i>QTI-P</i> 31 ítems, Nivel 1, 8 factores 1 ^{er} orden, con Covarianzas	671.9	406	<.001	1.65	.941	.949	.949	.046	.042	-140.1
M3	Estructura Circunfleja, 31 ítems, Nivel 1	867.4	432	<.001	2.01	.910	.916	.917	.074	.052	3.4
M4	Estructura Circunfleja, 31 ítems, Nivel 1 Factores 1 ^{er} y 2 ^{do} orden, Nivel 2	1180.0	850	<.001	1.39	.909	.917	.919	.194	.046	-519.9
M5	Estructura Circunfleja, 31 ítems, Nivel 1 Multinivel contextual, Nivel 2	1137.8	1114	.303	1.02	.994	.994	.995	.069	.011	-1090.1
M6	Estructura Circunfleja, 31 ítems, Nivel 1 Multinivel contextual, Nivel 2, Rendimiento académico, Niveles 1 y 2	1523.8	1570	.794	.97	1.009	1.000	1.008	.088	.000	-1616.2

Nota: Chi-cuadrado (χ^2), grados de libertad (gl), probabilidad de χ^2 [*p*(χ^2)], Bentler-Bonett Non Normed Index (NNFI), Comparative Fit Index (CFI), Bollen Fit Index (IFI), Standardized Root Mean-Square Residual (SRMR), Root Mean-Square Error of Aproximation (RMSEA) y Akaike Information Criterion (AIC).

En resumen, ahora el *QTI-P* en castellano queda con 31 ítems (se presentan en el Anexo con un *), por lo que hay tres escalas que han quedado con 3 variables observables (Amable, Insatisfecha e Impositiva), otras tres que han quedado determinadas por 4 variables (Directiva, Acomodatícia y Represiva), y dos que han quedado con 5 variables (Comprensiva e Insegura). Este recorte en el número de variables por escala no supone un problema, pues en un *AFC* con más de un factor (escalas), se considera suficiente para la métrica sea correcta que cada factor tenga tres variables (Kenny y McCoach, 2003). Los valores α obtenidos son buenos en todas las escalas, con un rango entre .63, escala Inseguridad, y .87, escala Represiva (ver Tabla 2).

(H2) El ajuste del modelo circunflejo (M3) se adecúa a los datos (Tabla 3), cumpliéndose la hipótesis del *MITB*, puesto que todos los indicadores son buenos. Las diferencias en *AIC* respecto del M2 son favorables al M2, pero aceptamos el M3 como mejor representación porque los indicadores de ajuste son buenos y responde a las hipótesis del *MITB*.

no ajusta, pues, aunque el *RMSEA* es bueno y la razón χ^2 /gl también es aceptable, los otros indicadores no lo son (Tabla 3). Por lo tanto, se ha depurado el modelo completo, quitando los ítems que disminuían la fiabilidad de la escala (ítems 16, 37 y 41), con λ inferiores a .30 (ítem 30) o que correlacionaban más alto con los ítems de otra escala (ítems 6, 10, 12, 14, 15, 19, 23, 24, 25, 34, 36, 40, y 42) hasta llegar a los 31 ítems que participan en el Modelo 2 (M2). Este modelo tiene un buen ajuste en todos los índices, además las varianzas de las escalas y las covarianzas entre escalas son estadísticamente significativas, así como todos los efectos de las escalas sobre los correspondientes ítems (λ). Las diferencias de *AIC* entre M1 y M2 son mucho más grandes del valor de 7 a favor de M2, por lo que aceptamos M2 como la mejor representación métrica de los datos del cuestionario *QTI-P*. No se exponen los resultados detallados de M2, pero los valores obtenidos de los efectos de las escalas sobre los ítems del *QTI-P* son sensiblemente iguales a los presentados en la Figura 3 (M6).

(H3) Las *ICC* de las escalas son muy altos (Tabla 2). Si exceptuamos la *ICC* de la escala Insegura (.10), las otras se sitúan en un rango entre .24 (Comprensiva) y .47 (Represiva). En ambientes escolares suelen ser infrecuentes *ICC* superiores a .30 (Lüdtke et al., 2008). Las *ICC* de las dimensiones son también muy altos, .35 en Control y .42 en Afiliación. Estos valores confirman la influencia del ambiente aula en las percepciones e invitan a realizar *AFC* multinivel (M4 y M5), téngase en cuenta que el *ICC* indica el porcentaje de similitud en la variable medida dentro de la clase en comparación con las otras clases, a mayor *ICC* mayor parecido de los alumnos dentro de la clase. M4 ajusta bien a los datos, además, todos los coeficientes son significativos y con signos coherentes respecto de los valores obtenidos, pues hay coherencia métrica variable-escala, el valor del *AIC* es mucho menor que el de M2 y el de M3, por lo que M4 puede ser una buena representación de los datos obtenidos. Cuando en el nivel 2 se usan las medias de las puntuaciones de *QTI-P* para cada escala dentro de cada clase (M5), los indicadores

de ajuste de conjunto son muy buenos (Tabla 3), y todos los parámetros han resultado ser significativos y coherentes en lo que a los signos se refiere. La lectura de los efectos de nivel 2 (la media de cada respectiva escala por clase), indica que, a más elevada media de clase en una escala, más puntuación tendrá un niño de esa clase en un ítem de esa escala. Las diferencias entre los índices de ajuste, incluido el *AIC*, con los modelos anteriores son siempre mejores en M5, por tanto, aceptamos este modelo como la mejor representación de ajuste de los resultados obtenidos del *QTI-P*. Por otra parte, se observa que los efectos de las escalas sobre los ítems y de las dimensiones sobre las escalas apenas cambian en M5 respecto a M2, M3 y M4, lo cual es un indicador de la estabilidad del modelo.

Así, de acuerdo con los valores de la Figura 3 llevados al M5, la ecuación de pronóstico para una niña *i* en una clase *j*, para el ítem 26 (*QTI26_{ij}*, perteneciente a la escala Amable, *Amable_i*), y a la variable contextual Amable, *ContextualAmable_j*, su valor pronosticado será:

$$QTI26'_{ij} = 1.332 * Amable_i + .898 * ContextualAmable_j \quad (2)$$

Siendo, en la Ecuación 2, el valor 1.332 el efecto de la puntuación factorial de la niña *i* sobre *QTI26*, que es el mismo para todos los niños para pronosticar cada valor de *QTI26*, como en el análisis factorial exploratorio; *Amable_i* es la puntuación factorial de la niña *i* en la escala Amable, y es un valor propio de cada niño en esta escala; .898 es el valor del coeficiente de nivel 2 de la variable contextual Amable, y es un valor común para todos los niños; *ContextualAmable_j* es la media de la clase *j* en las variables de la Escala Amable, y es propia de cada clase, al haber 19 aulas habrá 19 valores distintos; y, por último, no se ha incluido *E_{26,ij}*, porque sería el error de pronóstico del niño *i* en la clase *j* para el pronóstico de *QTI26* (*QTI26'_{ij}*). En la Ecuación 2, el término $1.332 * Amable_i$ es de primer nivel (propio para cada niña), mientras el término $.898 * ContextualAmable_j$ es de nivel 2 (específico para cada clase).

En la Ecuación 2, el término de nivel 2 ($.898 * ContextualAmable_j$) funciona como un intercepto, pues si se toma una clase (*j*) determinada, por ejemplo, para la clase 122, su valor de *ContextualAmable* para todos los niños de esa clase es de .692, por lo que para todos los niños de la clase el valor fijo es de .621 ($.898 * .692$), que se puede considerar como el nivel inicial común (o intercepto) para todos los niños de esta clase.

(H4) La Tabla 3 muestra que los resultados de conjunto de M6 son muy buenos estadísticamente y en la Figura 3 se representa gráficamente. Se observa cómo Rendimiento Académico es un factor único muy estable (con varianza significativa, $p < .001$), que influye significativamente en cada calificación académica: Matemáticas, Ciencias, Castellano y Valenciano. En el nivel 1, estudiante, sólo la dimensión Afiliación influye de manera positiva y significativa sobre el factor Rendimiento Académico. En el nivel 2, Aula, los componentes del *QTI-P* no influyen en el Rendimiento Académico;

pero Rendimiento Académico forma un factor a nivel 2 que determina la calificación media de cada clase en las calificaciones académicas, y cada una de ellas condiciona a su vez la puntuación individual de cada alumna en cada materia. De este modo, la calificación final de una alumna *i* (nivel 1) perteneciente a la clase *j* (nivel 2) en Ciencias sería (ver Figura 3):

$$Ciencias'_{ij} = 1.059 * RendimientoAcadémico_i + .996 * ContextualCiencias_j \quad (3)$$

Los términos de la Ecuación 3 son similares a los de la Ecuación 2, no se incluye el error de pronóstico en las Ecuaciones 3 y 5, porque es el error de pronóstico de la calificación en Ciencias (*Ciencias'_{ij}*). Téngase en cuenta que a su vez la variable *ContextualCiencias_j* es generada mediante los valores de nivel 2:

$$ContextualCiencias'_j = 1.023 * RendimientoAcadémico_j \quad (4)$$

O lo que es lo mismo, substituyendo el componente determinístico de la Ecuación 4 en la 3, para pronosticar la puntuación de un alumno en Ciencias:

$$Ciencias'_{ij} = 1.059 * RendimientoAcadémico_i + .996 * (1.023 * RendimientoAcadémico_j) \quad (5)$$

De la misma manera que en la Ecuación 2, la Ecuación 5 indica que la calificación de nivel 2 de cada niño en la clase *j* es común para todos los otros niños de esa clase, por lo que el valor de *RendimientoAcadémico_j* (valor de nivel 2, multiplicado por sus correspondientes coeficientes) actúa como un intercepto para los alumnos de la clase *j*.

Discusión

Este estudio confirma la fiabilidad y validez estructural de la versión española de *QTI* para alumnos de cursos superiores de educación primaria (*QTI-P*). *QTI-P* mantiene 31 ítems de los 48 originales distribuidos en las ocho escalas, todas conservan 3 o más ítems, y presentan mejores índices de fiabilidad, porcentajes de varianza explicada y capacidad para diferenciar entre clases, por los elevados valores de *ICC*, que los obtenidos en otros estudios (Goh y Fraser, 1998; den Brok et al., 2005; Kokkinos et al., 2009; Scott y Fisher, 2004; Telli y den Brok, 2012), muy especialmente en las escalas Insegura, Impositiva y Acomodatícia, todas ellas más relacionadas con la dimensión Control.

Quizá estos resultados pueden ser explicados por algunas de las novedades incorporadas en el estudio como son: (a) el empleo de una escala de respuesta de 5 puntos, que posibilita que exista suficiente diferenciación (Telli y den Brok, 2012); (b) la utilización de la conducta del maestro tutor como referente del ambiente, más representativa de la cultura en las aulas de primaria que la de un profesor de una asignatura (Maulana et al., 2012); y (c) el énfasis en considerar tanto la

contribución de cada ítem a su escala teórica como su aportación a las relaciones entre escalas (Wubbels et al., 2012).

Por lo que se refiere a la validación de la estructura circunfleja (Hipótesis 2, Modelo 3), los resultados muestran que, usando puntuaciones individuales (alumno) de las variables observables, las 8 escalas pueden ser explicadas por dos dimensiones independientes, Control y Afiliación, y que ocupan posiciones fijas en la estructura circular de acuerdo con lo pronosticado por el modelo *MITB*. Otros intentos no han resultado exitosos (Telli y den Brok, 2012), o su confirmación se ha realizado de forma aproximada mediante la comparación entre patrones de correlaciones entre escalas (Telli y den Brok, 2012). Por último, al igual que en la mayoría de los estudios, la dimensión Afiliación es la que más apoyo empírico recibe (Fisher et al., 2011), pues en nuestro caso tiene mejor *VME* que Control, buena *ICC* y presenta efectos significativos sobre el factor Rendimiento Académico a nivel individual. Al mismo tiempo, en la versión española también se refuerza la dimensión Control, puesto que contribuye en mayor grado a la percepción de los estudiantes de las *IMA* (media = 0.56) y los alumnos dentro de las clases alcanzan mayor similitud en esta dimensión ($ICC = .352$) que en estudios anteriores.

Por lo que se refiere a la hipótesis 3 (Modelo 5), las *ICC* obtenidas son muy altas, tanto en las 8 escalas como en las dos dimensiones, mostrando diferencias entre aulas a la vez que similitud entre los estudiantes de la misma aula en su percepción de la conducta de la maestra tutora (Lüdtke et al., 2008). Se trata del primer estudio con el *QTI-P* que ha usado un modelo multinivel, con puntuaciones individuales (alumno) de las variables observables explicadas por las escalas y éstas por las dimensiones bajo las hipótesis del *MITB* en el primer nivel, y con puntuaciones agregadas (aula) de las escalas explicando las variables observables en el segundo nivel. Previamente, solo el estudio de Fisher et al. (2011) había confirmado la estructura multinivel del *QTI-P*, pero usando las puntuaciones agregadas de aula de cada escala tanto en el primer y en el segundo nivel, y sin incluir en el modelo los efectos de las escalas sobre los ítems. En nuestra investigación se comprueba que, en la puntuación final de cada niño en cualquier ítem del *QTI-P*, influye la puntuación media de la escala a la que pertenece dicho ítem en su aula, más la puntuación específica del niño en ese ítem. De este modo, la media de la escala actúa como un intercepto común para los niños de una misma clase.

Por lo que se refiere a la hipótesis 4 (Modelo 6), sobre predicción del rendimiento, se ha comprobado que las percepciones del alumnado sobre la conducta del profesorado no dependen del rendimiento académico. En cambio, la calificación final obtenida por cada alumno depende de la percepción de la conducta afiliativa del maestro a nivel individual y de la calificación media agregada de cada asignatura a nivel de aula (ver Figura 3 y Ecuaciones 3 a 5).

Por último, ¿qué percepción tienen los estudiantes españoles de la conducta interpersonal de sus maestros con ellos? Las puntuaciones de los estudiantes españoles en todas las

escalas del *QTI-P* son más bajas que en los estudios previos, lo que puede dar la impresión de que el alumnado percibe que el profesorado se ocupa poco de ellos. Una observación más detenida señala que las medias son particularmente bajas en las escalas que cargan negativamente en la dimensión de Afiliación, muy especialmente las que comparten Control bajo, conductas Inseguras e Insatisfechas. Esto es, que los estudiantes españoles, aunque perciben menos conductas Cooperativas y de Dominancia en sus maestros que los estudiantes de otros países, perciben todavía menos conductas de Oposición y de bajo Control. Ello se concreta en que las conductas docentes que combinan Oposición y bajo Control son muy infrecuentes en nuestro país y, sobre todo, en que la Afiliación resulte muy alta y el Control moderado, muy superior y ligeramente superior a lo que ocurre en otros países, respectivamente, de acuerdo con las indicaciones propuestas por den Brok et al. (2006), lo que resulta en una imagen positiva del profesorado.

Estos resultados concuerdan con los dos únicos trabajos que hemos encontrado en los que participan estudiantes de origen hispano: a) En el estudio multicultural de den Brok, Levy, Wubbels y Rodríguez (2003, citado en den Brok y van Tartwijk, 2015) se encontró que los estudiantes de secundaria hispanoamericanos generalmente percibían que las *IMA* era más afiliativas y dominantes que los otros alumnos de la muestra. b) Los estudiantes españoles de primer y segundo curso puntuaron más alto en ambas dimensiones (García Bacete et al., 2014) que en otro estudio realizado con niños de los países bajos de la misma edad (Zijlstra et al., 2013).

En conclusión, hemos visto que el *QTI-P*, al solicitar a los alumnos percepciones compartidas sobre la maestra, más que individuales, permite describir un ambiente propio en cada aula, entendido como la interacción de la maestra con el alumnado como colectivo, y que el *QTI-P* lo mide de modo fiable. En esta afirmación destacamos tres aspectos: (a) la idoneidad del *QTI-P* para captar el ambiente único de cada aula; (b) la relevancia del aula como grupo de alumnos; y (c) la relevancia de la maestra como constructora del ambiente. El efecto del ambiente es tal que tiene una clara influencia en el rendimiento académico individual, junto con el rendimiento medio del aula, sin que la percepción del ambiente del aula se vea influida por el rendimiento académico.

En definitiva, las investigadoras y maestras españolas tienen disponibles dos versiones del *QTI* que les permite evaluar la interacción profesora-alumnado a lo largo de toda la educación primaria, la versión *QTI-EP*, válida para los escolares de 1º a 3º de educación primaria, y la versión *QTI-P* para los de 4º a 6º; un inconveniente de nuestra muestra es que no ha podido contar con alumnos de 5º curso por limitación de presupuesto, pero cabe esperar que si el modelo propuesto se ajusta a los datos de nuestros alumnos de 4º y 6º, también se ajustará a los de 5º. Ahora el reto consiste en determinar cómo el profesorado puede crear tales relaciones positivas (Wubbels et al., 2012). En nuestra opinión, y de acuerdo con las conclusiones enumeradas, dos ejes deben dirigir estas intervenciones: (a) la necesaria autorreflexión del profe-

sorado sobre su conducta interactiva en cada aula, a partir de las respuestas proporcionadas por el alumnado y de sus propias respuestas al *QTI-P* (Brekelmans et al., 2011); (b) el diseño de procesos de acompañamiento y consultoría al profesorado, que le facilite las condiciones para desarrollar y aplicar intervenciones curriculares y universales de gestión socioemocional dirigidas a todo el alumnado del aula (García Bacete et al, 2019). Finalmente, como líneas de futuro señalamos la realización de estudios que incorporen un mayor

número de aulas, el análisis de los perfiles conductuales de profesorado de primaria y la extensión de la validación del *QTI* a todos los cursos de la educación secundaria en España.

Agradecimientos.- Este trabajo ha sido posible a las ayudas concedidas al primer autor por el Ministerio de Economía y Competitividad del Gobierno de España (EDU2012-35930) y por la Universidad Jaume I (UJI-B2019-29).

Referencias

- Abelló, D., Alonso-Tapia, J., & Panadero, E. (2020). Development and validation of the Teaching Styles Inventory for Higher Education. *Anales de Psicología*, *36*(1), 143-154. doi: 10.6018/analesps.370661
- Bentler, P. M. (2016). *EQS v 6.3 structural equations program*. Encino, CA: Multivariate Software.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research*, *33*, 261-304. doi: 10.1177/0049124104268644
- Brekelmans, M., Mainhard, T., den Brok, P., & Wubbels, T. (2011). Teacher control and affiliation: Do students and teachers agree? *Journal of Classroom Interaction*, *46*, 17-26.
- Charalampous, K., & Kokkinos, C.M. (2017). The structure of pre-adolescents' perceptions of their teacher's interpersonal behaviours and their relation to pre-adolescents' learning outcomes. *Educational Studies*, *44*(2), 167-189. doi:10.1080/03055698.2017.1347492
- Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research*, *77*(1), 113-143. doi: 10.3102/003465430298563
- den Brok, P., Brekelmans, M., & Wubbels, T. (2006). Multilevel issues in research using student' perceptions of learning environments: The case of the Questionnaire on Teacher Interaction. *Learning Environment Research*, *9*, 199-213. doi: 10.1007/s10984-006-9013-9
- den Brok, P., Fisher, D., & Scott, R. (2005). The importance of teacher interpersonal behaviour for student attitudes in Brunei primary science classes. *International Journal of Science Education*, *27*, 765-779. doi: 10.1080/09500690500038488
- den Brok, P., & van Tartwijk, J. (2015). Teacher-student interpersonal communication in international education. In M. Hayden, J. Levy, & J. J. Thompson, *The SAGE Handbook of research in international education* (pp. 309-324). City Road, London: SAGE Publications Ltd. doi: 10.4135/9781473943506.n21
- Fisher, D., Henderson, D., & Fraser, B. (1995). Interpersonal behavior in senior high school biology classes. *Research in Science Education*, *25*, 125-133.
- Fisher, D., den Brok, P., Waldrip, B., & Dorman, J. (2011). Interpersonal behavior styles of primary education teachers during science lessons. *Learning Environment Research*, *14*, 187-204. doi: 10.1007/s10984-011-9093-z
- Fraser, B. J., & Walberg, H. J. (2005). Research on teacher-student relationships and learning environments: Context, retrospect and prospect. *International Journal of Educational Research*, *43*, 103-109. doi: 10.1016/j.ijer.2006.03.001
- García Bacete, F. J., Ferrà, P., Monjas, M. I., & Marande, G. (2014). Teacher-Students Relationships in First and Second Grade Classrooms. Adaptation of the Questionnaire on Teacher Interaction-Early Primary (*QTI-EP*). *Revista de Psicodidáctica*, *19*(1), 211-231. doi: 10.1387/RevPsicodidact.9081
- García Bacete, F. J., Jiménez, I., Muñoz-Tinoco, V., Marande, G., Monjas, M. I., Sureda, I., & cols. (2014). *El rechazo entre iguales en su contexto interpersonal: Una investigación con niños y niñas de primer ciclo de primaria*. Castellón: Fundación Dávalos-Fletcher
- García Bacete, F. J., Marande, G., & Mikami, A. Y. (2019). Changes in social climate after a two-year intervention for reducing peer rejection in early elementary education. *Journal of School Psychology*, *77*, 124-138. doi: 10.1016/j.jsp.2019.09.001
- Goh, S. C., & Fraser, B. J. (1998). Teacher interpersonal behaviour, classroom environment and student outcomes in primary mathematics in Singapore. *Learning Environments Research*, *1*, 199-229. doi: 10.1023/A:1009910017400
- Hox, J. J., Moerbeek, M., & Schoot, R. V. D. (2017). *Multilevel analysis. Techniques and applications* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Iacobucci, D. (2009). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology*, *20*(1), 90-98. doi: 10.1016/j.jcps.2009.09.003
- IBM Corp. (2016). *IBM SPSS statistics for Windows, Version 24.0*. Armonk, NY: IBM Corp.
- IBM Corp.
- Jimerson, S. R., & Haddock, A. D. (2015). Understanding the importance of teachers in facilitating student success: Contemporary science, practice, and policy. *School Psychology Quarterly*, *30*(4), 488-493. doi:10.1037/spq0000134
- Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling*, *10*(3), 333-351. doi: 10.1207/S15328007SEM1003_1
- Kokkinos, C. M., Charalampous, K., & Davazoglou, A. (2009). Interpersonal teacher behaviour in primary school classrooms: A cross-cultural validation of a Greek translation of the Questionnaire on Teacher Interaction. *Learning Environment Research*, *12*, 101-114. doi: 10.1007/s10984-009-9056-9
- Leary, T. (1957). *An interpersonal diagnosis of personality*. New York: Ronald.
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, *13*(3), 203-229. doi: 10.1037/a0012869
- Maulana, R., Opendakker, M-Ch., den Brok, P., & Bosker, R. (2012). Teacher-student interpersonal relationships during the first year of secondary. A multilevel growth curve analysis. In T. Wubbels, P. den Brok, J. van Tartwijk, & J. Levy (Eds.), *Interpersonal Relationships in Education: An Overview of Contemporary Research* (pp. 207-224). Rotterdam: Sense publishers. doi: 10.1007/978-94-6091-939-8_13
- Mindrila, D. (Ed.) (2017). *Exploratory factor analysis: Applications in school improvement research*. Hauppauge, NY: Nova Science.
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén and Muthén.
- Passini, S., Molinari, L., & Speltini, G. (2015). A validation of the Questionnaire on Teacher Interaction in Italian secondary school students: the effect of positive relations on motivation and academic achievement. *Social Psychology of Education*, *18*, 547-559. doi: 10.1007/s11218-015-9300-3
- Roorda, D. L., Koomen, H., Spilt, J. T., & Oort, F. J. (2011). The Influence of Affective Teacher-Student Relationships on Students' School Engagement and Achievement: A Meta-Analytic Approach. *Review of Educational Research*, *81*(4), 493-529. doi: 10.3102/0034654311421793
- Scott, R. H., & Fisher, D. L. (2004). Development, validation and application of a Malay translation of an elementary version of the Questionnaire on Teacher Interaction. *Research in Science Education*, *34*, 173-194. doi: 10.1023/B:RISE.0000033759.09807.50
- Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences*, *42*(5), 893-98. doi: 10.1016/j.paid.2006.09.017
- Telli, S., & den Brok, P. (2012). Teacher-student interpersonal behaviour in the Turkish primary to higher education context. In T. Wubbels, P. den Brok, J. van Tartwijk, & J. Levy (Eds.), *Interpersonal relationships in education: An overview of contemporary research* (pp. 187-206). Rotterdam: Sense publishers. doi: 10.1007/978-94-6091-939-8_12
- Tennant, J. E., Demaray, M. K., Malecki, Ch. K., Terry, M. N., Clary, M., & Elzinga, N. (2015). Students' ratings of teacher support and academic and social-emotional well-being. *School Psychology Quarterly*, *30*(4), 494-512. doi: 10.1037/spq0000106
- Wubbels T., & Brekelmans, M. (2005). Two decades of research on teacher-student relationships in class. *International Journal of Educational Research*, *43*, 6-24. doi: 10.1016/j.ijer.2006.03.003
- Wubbels, T., Brekelmans, M., den Brok, P., Levy, J., Minhard, T., & van Tartwijk, J. (2012). Let's make things better. Developments in research on interpersonal relationships in education. In T. Wubbels, P. den Brok, J. van Tartwijk & J. Levy (Eds.), *Interpersonal relationships in education: An overview of contemporary research* (pp. 225-249). Rotterdam: Sense publishers. doi: 10.1007/978-94-6091-939-8_14

Wubbels, T., Créton, H. A., & Hooymayers, H. P. (1985). *Discipline problems of beginning teachers, interactional behavior mapped out*. Paper presented at the American Educational Research Association Annual Meeting, Chicago.

Wubbels, T., & Levy, J. (1991). A comparison of interpersonal behaviour of Dutch and American teachers. *International Journal of Intercultural Relationships*,

15, 1-18.

Zijlstra, A., Wubbels, T., Brekelmans, M., & Koomen, H. (2013). Child perceptions of teacher interpersonal behavior and associations with mathematics achievement in Dutch early grade classrooms. *The Elementary School Journal*, 113(4), 517-540. doi:10.1086/669618

Anexo

Cuestionario Interacciones entre el Maestro y el Alumnado en un Aula de Primaria (*QTI-P*, *Questionnaire on Teacher Interaction – Primary Education*).

Nº Ítem	Enunciado Ítem	Escala
1*	Todos le prestamos atención a este maestro.	Directiva
2*	Este maestro es simpático.	Amable
3*	Este maestro confía en nosotros.	Comprensiva
4*	Este maestro nos permite trabajar en cosas que nos gustan.	Acomodatícia
5*	Este maestro parece inseguro.	Insegura
6	Este maestro parece triste.	Insatisfecha
7*	Este maestro se enfada rápidamente.	Represiva
8*	Este maestro nos hace trabajar duro.	Impositiva
9*	Aprendemos mucho con este maestro.	Directiva
10	A este maestro le gusta reírse.	Amable
11*	Este maestro se da cuenta cuando no comprendemos alguna cosa.	Comprensiva
12	Los alumnos podemos decidir algunas cosas en la clase de este maestro.	Acomodatícia
13*	Este maestro tiene poca confianza en sí mismo.	Insegura
14	Este maestro está de mal humor.	Insatisfecha
15	Este maestro nos menosprecia.	Represiva
16	En las clases de este maestro tenemos que estar callados.	Impositiva
17*	Este maestro capta nuestra atención.	Directiva
18*	Las clases de este maestro son agradables.	Amable
19	Cuando no entendemos algo este maestro nos lo explica otra vez.	Comprensiva
20*	Este maestro nos deja mucho tiempo libre en clase.	Acomodatícia
21*	Este maestro es un poco tímido, vergonzoso.	Insegura
22*	Este maestro piensa que no sabemos hacer las cosas bien.	Insatisfecha
23	Este maestro se burla de nosotros.	Represiva
24	Los exámenes de este maestro son difíciles.	Impositiva
25	Este maestro sabe todo lo que pasa en esta clase.	Directiva
26*	Nos gusta este maestro.	Amable
27*	Este maestro presta atención a lo que decimos.	Comprensiva
28*	Este maestro nos permite elegir con quien queremos trabajar.	Acomodatícia
29*	Este maestro no sabe qué hacer cuando perdemos el tiempo en nuestras cosas.	Insegura
30	Este maestro piensa que los alumnos nos copiamos.	Insatisfecha
31*	Este maestro nos grita.	Represiva
32*	Este maestro es exigente cuando corrige nuestras tareas y exámenes.	Impositiva
33*	Este maestro explica las cosas con claridad.	Directiva
34	Este maestro nos ayuda con nuestro trabajo.	Amable
35*	Este maestro sabe cómo nos sentimos	Comprensiva
36	Este maestro nos deja que nos entretengamos en nuestras cosas.	Acomodatícia
37	Este maestro nos deja que le digamos lo que tiene que hacer.	Insegura
38*	Este maestro piensa que no sabemos nada.	Insatisfecha
39*	Este maestro se enfada por cualquier cosa.	Represiva
40	Este maestro nos da un poco de miedo.	Impositiva
41	El maestro tiene claro lo que quiere que ocurra en clase	Directiva
42	Este maestro se interesa por nosotros.	Amable
43*	Este maestro nos escucha.	Comprensiva
44*	Este maestro nos permite elegir en qué queremos trabajar.	Acomodatícia
45*	Este maestro actúa como si no supiera qué tiene que hacer.	Insegura
46*	Este maestro nos amenaza con castigarnos.	Insatisfecha
47*	Este maestro tiene mal carácter.	Represiva
48*	Este maestro es severo.	Impositiva

*Los 31 ítems que ajustan en la versión en castellano.

La expresión "Este maestro" se sustituye por el nombre del maestro o la maestra de la clase correspondiente.