University of Groningen

Pupils' perceptions of teaching behaviour

Maulana, Ridwan; Helms-Lorenz, Michelle; van de Grift, Wim

Published in:
International Journal of Educational Research

DOI:
[10.1016/j.ijer.2014.11.002](https://doi.org/10.1016/j.ijer.2014.11.002)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Maulana, R., Helms-Lorenz, M., & van de Grift, W. (2015). Pupils' perceptions of teaching behaviour: Evaluation of an instrument and importance for academic motivation in Indonesian secondary education. *International Journal of Educational Research*, 69, 98-112. <https://doi.org/10.1016/j.ijer.2014.11.002>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Pupils' perceptions of teaching behaviour: Evaluation of an instrument and importance for academic motivation in Indonesian secondary education

Ridwan Maulana*, Michelle Helms-Lorenz, Wim van de Grift

Department of Teacher Education, Faculty of Social and Behavioural Sciences, University of Groningen, The Netherlands

ARTICLE INFO

Article history:

Received 18 August 2014

Received in revised form 5 November 2014

Accepted 24 November 2014

Available online 23 December 2014

Keywords:

Teaching behaviour
Academic motivation
Rasch model
Multilevel model
Secondary education

ABSTRACT

The present study investigates the psychometric quality of a measure tapping pupils' perceptions of teachers' teaching behaviour in the Indonesian context. It also examines the relationship between pupils' perceptions of teaching behaviour and their perceived academic motivation. Surveys from a total of 4000 pupils of 200 teachers/classes were analysed using classical test analyses and Rasch modelling. Results reveal that the Indonesian version of the instrument is reliable and valid for measuring teaching behaviour in the Indonesian context. Teachers' teaching behaviour is a significant predictor of pupils' controlled and autonomous motivation. The effect of teaching behaviour is stronger for pupils' autonomous motivation than for controlled motivation. Implications of findings for research and educational practices were discussed.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Putting effort into enhancing the quality of education is an ongoing educational concern all over the world. Teachers, who are seen as central figures in promoting pupils' academic performance up to the highest level possible, have been a major focus of current educational research. The central role of the teacher is clearly highlighted in the teacher effectiveness literature revealing that teacher factors explain higher proportions of variance in the success of pupils' learning than other factors (Creemers, 1994; Kyriakides, Creemers, & Antoniou, 2009). Among other factors, teaching behaviour has been regarded as one of the most prominent factors affecting pupils' learning and achievement (Creemers, 1994; Hattie, 2012). To improve educational outcomes, developing and sustaining effective teaching behaviour should be a priority for schools.

Although research recognizes the importance of teaching behaviour for pupils' outcomes, several limitations remain. One problem we are currently facing is the development of guidelines for evidence-based support to develop and sustain (amongst other objectives) the improvement of effective teaching behaviour. Particularly, this guideline is critically important for pre-service and beginning teachers during the first years of teaching. This is because past research revealed that inexperienced teachers generally face difficulties in performing effective teaching behaviour (Van de Grift, 2007; Van de Grift & Helms-Lorenz, 2012; Van de Grift, Helms-Lorenz, & Maulana, 2014). Guidelines focusing on the development of effective teaching behaviour targeting at the zone of proximal development will be beneficial for (inexperienced) teachers. Another issue is that there is a gap between Western and Asian countries regarding the advancement and access of

* Corresponding author at: Department of Teacher Education, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands.
E-mail address: r.maulana@rug.nl (R. Maulana).

knowledge. In Western countries, attempts to enhance effective teaching behaviour have been more rapid and more well disseminated in the scientific outlet. In the Netherlands, for example, the beginning teacher professional development program (i.e., induction arrangement) has been implemented as a national program for supporting effective teaching behaviour (Helms-Lorenz, Slof, & Van de Grift, 2013). In Asian countries like Indonesia such a program remains a topic of discourse. Attempts for supporting professional development of (inexperienced) teachers in the Asian region is therefore needed.

Finally, many countries strive to motivate and engage their pupils to participate in learning activities. In secondary education, particularly, there is evidence that pupils' motivation tends to decrease over time (Maulana, Opdenakker, & Bosker, 2013; Opdenakker, Maulana, & den Brok, 2012). Research originated from Western countries showed that teaching behaviour has a powerful impact on sustaining pupils' motivation (den Brok, Fisher, & Scott, 2005; Opdenakker et al., 2012; Wubbels & Brekelmans, 2005). However, little is known about the relationship between teaching behaviour and Asian pupils' motivation to learn. Particularly, past research in the Dutch context revealed that effective teaching behaviour, measured in terms of teacher-pupil interpersonal relationships, affects pupils' intrinsic motivation positively (Opdenakker et al., 2012). However, this study did not find a significant effect of teaching behaviour on pupils' extrinsic motivation. Considering that Western and Asian contexts view the importance of intrinsic and extrinsic motivation in a different way (Leung, 2001), little is known about the effect of teaching behaviour on pupils' motivation found in the Western context in comparison to the Asian context.

The current study aims to fill the voids in the knowledge base by presenting findings from the Indonesian context in two ways. First, we examine the usefulness of an instrument previously developed in the Western context that can serve as a practical guideline for supporting professional development of (inexperienced) teachers targeting at their zone of proximal development. Additionally, the study aims to shed more light on the inconclusiveness of teaching behaviour effects on pupils' motivation in the Indonesian context.

2. Theoretical frameworks

2.1. Teaching behaviour

Research on educational effectiveness and learning environments, which serves as a theoretical framework for studying teaching behaviour, has shown that classroom processes are important predictors of pupils' learning (Creemers & Kyriakides, 2008). In contemporary conceptualizations of teaching and learning processes, both cognitive and motivational aspects are taken into account. Teaching behaviour is a central element in the classroom processes (Creemers, 1994; Hattie, 2012). Based on reviews of evidence-based effective teaching behaviour research, six observable domains of teaching behaviour were identified that impact on pupils' learning and outcomes (Van de Grift, 2007, 2013). These domains are safe and stimulating learning climate, efficient classroom management, clarity of instruction, activating learning, differentiation, and teaching learning Strategies. These six domains are in accordance with other models and empirical findings (Danielson, 2013; Pianta & Hamre, 2009).

Safe and stimulating learning climates refers to aspects of teacher-pupil relationships such as creating a relaxing learning atmosphere, showing respect to pupils and ensuring that pupils respect the teacher and their peers, and encouraging self-confidence of pupils (Cornelius-White, 2007; Hattie & Clinton, 2008; Opdenakker et al., 2012; Smith, Baker, Hattie, & Bond, 2008; Teodorovic, 2011; Willms & Somers, 2001). Effective classroom management is another important domain of teaching behaviour. Research showed that efficient classroom management is an important predictor of pupils' learning and outcomes (Carnine, Dixon, & Silbert, 1998; Houtveen, Booij, De Jong, & Van de Grift, 1999; Maulana, Opdenakker, Stroet, & Bosker, 2012; Scheerens & Bosker, 1997). Several indicators of teachers' behaviour associated with efficient classroom management include ensuring that the lesson begins and ends on time, managing lesson transition efficiently, minimizing time for task-unrelated matters, dealing with pupils' misbehaviour efficiently, preparing the lesson well and displaying good lesson structure (Creemers, 1994; Marzano, 2003; Maulana et al., 2012; Opdenakker & Minnaert, 2011; Wang, Reynolds, & Walberg, 1995; Yair, 2000). Clarity of instruction covers aspects of instructional behaviour that are important for pupils' learning such as clear lesson structure, good interchange of explanations and lesson presentations, management of independent work, and clear assignment of individual and group works (Creemers, 1994; Kindsvatter, Willen, & Ishler, 1988; Mortimore, Sammons, Stoll, Lewis, & Ecob, 1988; Rosenshine, 1980). Additionally, checking that pupils understand the learning material is also important (Kame'enui & Carnine, 1998; Pearson & Fielding, 1991; Pearson & Gallagher, 1983; Hattie & Clinton, 2008; Rosenshine & Meister, 1997; Smith et al., 2008).

Activating learning integrates behaviour that optimizes learning outcomes such as promotion of active learning, intensification of instructions and avoiding excessive work seats (Hampton & Reiser, 2004; Lang & Kersting, 2007; Rosenshine & Berliner, 1978). Other important aspects of activating learning for productive learning outcomes are activation of prior knowledge, make use of "advance organizers", and make sure that pupils are aware of the relevance of the lesson content (Pressley et al., 1992; Nunes & Bryant, 1996). Activating learning is related to the quality of teacher-pupil and peer interactions (Meeuwisse, Severiens, & Born, 2010). Subsequently, the quality of teacher-pupil interactions is related to the improvement of motivational outcomes (Opdenakker et al., 2012; Maulana & Opdenakker, 2013). Differentiation in teaching means that teachers recognize the characteristics of pupils they teach and understand that different pupils have learning needs in order to progress in learning. Teachers should emphasize the needs of each individual student when displaying

instructional methods. In the literature of differentiation, the term “differentiated instruction” is recognized. This term refers to “a philosophy of teaching purporting that students learn best when their teachers effectively address variance in students’ readiness levels, interests, and learning profile preferences” (Tomlinson, 2005, p. 263). The main goal of differentiated instruction is to maximize the learning potential of each student (Tomlinson, 2003). Research showed that differentiated instruction leads to better learning performance (Koeze, 2007; Reis, McCoach, Little, Muller, & Kaniskan, 2011). Several indicators of teaching behaviour associated with differentiation in teaching include devoting extra time and additional instructions, pre-teaching and re-teaching, and implementing various effective teaching methods (Houtveen et al., 1999; Kindsvatter et al., 1988; Lundberg & Linnakylä, 1992; Pearson & Fielding, 1991; Sijtsma, 1997).

Finally, teaching learning strategies is “a heuristic that serves to support students, facilitating the development of internal procedures that enable them to perform the higher-level procedures (Van de Grift, 2007). Teaching learning strategy suggests that teachers support the use of metacognitive strategies in their classroom teaching. Metacognitive strategies provide a framework to help pupils achieve higher level of learning skills (Carnine et al., 1998). One way of supporting the use of metacognitive strategies is scaffolding. Scaffolding is a form of temporary support provided by teachers (or by peer pupils) that function as a bridge between pupils’ existing and desired skills. Research showed that teachers who display modelling explicitly, deliver scaffolding and provide corrective feedback contribute significantly to the performance of their pupils (Good & Brophy, 1989; Hattie & Clinton, 2008; Houtveen & Van de Grift, 2007; Rosenshine & Stevens, 1986; Slavin, 1996; Smith et al., 2008).

2.2. Pupil perceptions of teaching behaviour

In addition to observation measure, pupil perceptions of teaching behaviour is a frequently-used measure of classroom practices. Pupil perceptions of teachers’ behaviour contributes to the understanding of the quality of teaching practices for several reasons. First, pupil perceptions are based on their (classroom) experiences over time, not merely from a single snapshot of observations. This suggests that pupil perceptions may actually be more important than what outsiders would observe since pupil perceptions steer their learning behaviour, based on own judgements and insights. In line with this, research revealed that active pupils’ engagement in academic activities depends on their perceptions of the quality and integrity of teachers’ behaviour (den Brok, Brekelmans, Wubbels, 2004). Additionally, there is evidence that pupil perceptions are mostly more predictive of their learning outcomes compared to observations (De Jong & Westerhof, 2001; Seidel & Shavelson, 2007) and teachers’ self-report of their own teaching behaviour (Maulana, Opdenakker, den Brok, & Bosker, 2011; Scantlebury, Boone, Kahle, & Fraser, 2001). Practically, measuring teaching quality in terms of pupil ratings is mostly preferred because it is considered as the most cost-effective method of data collection compared to observations (Fraser, 1991).

2.3. Academic motivation: self-determination theory

Academic motivation covers aspects of motivational characteristics related to involvement in academic activities. In the Western context, self-determination theory has been widely used to study academic motivation. Although the theory posits that academic motivation is universally important for both Western and non-Western pupils’ learning outcomes (Chirkov & Ryan, 2001), the use of SDT for research in the Asian context is rather scarce.

Self-determination theory recognizes two types of motivation, namely intrinsic and extrinsic motivation. Intrinsic motivation is a prototype of self-determined activities, suggesting that involvement in certain tasks is based on volitional and autonomous reasons. Intrinsically motivated pupils involve in the academic activity for their own sake (non-instrumental). This can be contrasted from extrinsically motivated pupils who perform activities for achieving certain outcomes that are separable from the activity itself (instrumental). Hence, extrinsic motivation is driven by external pressures (Deci & Ryan, 1985). Extrinsically motivated behaviours can differ in the degree of the locus of causality or self-determination, although the behaviour performed is meant to achieve an instrumental end (Ryan & Deci, 2000).

The first is external regulation, which is considered as the least autonomous type of extrinsic motivation. When external regulation is at play, involvement in tasks is performed to meet an external demand such as getting compliments and avoiding punishments. The second is introjected regulation, which entails the exhibition of behaviours to please internal contingencies. In introjected regulation, one has introjected external demands. Behaviour is performed based on shame and guilt, suggesting that the form of behavioural style is experienced as more controlling. The third is identified regulation. This type of regulation suggests that behaviours are displayed as a reflection of awareness of valuing the regulation and the action is accepted as intrinsically important (Ryan & Deci, 2000; Vansteenkiste, Simons, Lens, Sheldon, & Deci, 2004). Empirically, external regulation is often combined with introjected regulation to form a *controlled motivation* composite (Vansteenkiste et al., 2004). Controlled motivation can be contrasted with an *autonomous motivation* composite containing identified regulation and intrinsic motivation, where regulations are regarded as originating from, and compatible with, the self. Autonomous motivation entails that the perceived locus of causality is highly internal and the behavioural regulations are experienced as highly self-determined (Niemic & Ryan, 2009).

Self-determination theory posits that autonomous motivation should be developed and sustained, while controlled motivation should be kept at a minimum level to optimize learning (outcomes). When a rather external regulation is at play, pupils will not learn as much as expected. Research in the Western context showed that intrinsic motivation is related to

task-oriented coping and investment in learning, whereas extrinsic motivation is associated with disengagement-oriented coping (Peetsma & Van der Veen, 2011; Thomson & Gaudreau, 2008). Intrinsic motivation has been shown to be an important predictor of academic outcomes such as achievement in school, career, and life in general (Gottfried, Gottfried, Cook, & Morris, 2005; Schick & Phillipson, 2009; Sternberg, 2001). Consistently, research from the Singaporean context revealed that intrinsic motivation is related to a positive view of learning (Chye, Liao, & Liu, 2012).

Controlled motivation is not valued in the Western context because it has detrimental effects on pupils' learning. Nevertheless, Asian researchers argued that both controlled and autonomous motivation contributes to East-Asian students' learning in an additive way (Leung, 2001; Zhu & Leung, 2010). Some cross-cultural scholars argued that autonomous motivation corresponds less with Eastern societies holding collectivistic values compared to Western societies holding individualistic values (Hofstede, 1991; Triandis, 1995). In a collectivist society, students' experience of autonomy is endorsed less because not only the self is important, but also the maintenance of social harmony and group support is highly valued and is regarded as a lifelong obligation (Olsen et al., 2002; Tseng, 2004). Within this social structure, students value learning not only because it is interesting for them, but also because fulfilling teachers' (and parents') expectations to study hard and to succeeding at school is personally important as well. Valuing controlled and autonomous motivation as complementary in non-Western contexts suggests that teaching behaviour may affect both types of pupils' academic motivation.

2.4. Teaching behaviour and academic motivation

Teaching behaviour is an important factor for motivational processes (Ryan & Deci, 2000; Kunter, Baumert, & Koller, 2007; Opdenakker et al., 2012). Studies in Western countries showed that teacher-pupil relationships are significant for pupils' individuation, personality and motivation development (Maulana et al., 2013; Opdenakker et al., 2012; Reeve, 2002; Stroet, Opdenakker, & Minnaert, 2013). Teachers' involvement serves as a facilitator of students' feelings of belongingness, which has a substantial impact on motivational outcomes and academic engagement (e.g., Maulana et al., 2012; Pavey, Greitemeyer & Sparks, 2011; Stroet et al., 2013; Wentzel, 1999). Pupils' perceptions of teachers' interpersonal closeness is positively and significantly related to pupils' autonomous motivation, but not to controlled motivation (Opdenakker et al., 2012). Experiences of connectedness, autonomy, and competence to teachers are shown to be associated with intrinsic motivation (Kunter et al., 2007). Efficient classroom management is related to positive development of subject-related interest (Kunter et al., 2007). Additionally, activating learning affects pupils' interest in learning (Fraser & Fisher, 1982; Kunter, 2005).

A study from the Indonesian context demonstrated that pupils' perceptions of teacher-pupils interpersonal relationships are positively and significantly related to pupils' controlled as well as autonomous motivation (Maulana et al., 2011, 2013). Based on studies reviewed above, there is an indication that the importance of teachers' teaching behaviour for academic motivation may differ depending on the form of motivation and on the contextual (i.e., Western vs. non-Western) background.

3. The current research: context and research questions

Based on research showing the importance of teaching behaviour for pupils' learning and outcomes reviewed earlier, as well as the fact that a robust diagnostic measure tapping pupils' perceptions of teaching behaviour for use in the Indonesian context is not present yet, it is our aim to focus the present research in this particular issue.

Culturally, the Indonesian society is collectivist in nature (Hofstede, 1991). This can be contrasted with most Western countries with the individualist culture. The classroom context of a collectivist culture reflects a social unit within the larger unit of hierarchical society. This hierarchical structure suggests that obedience to higher authority figures is valued. Furthermore, the Indonesian society is characterized by a very high power distance index, suggesting a high level of inequality of power within the society. Respect for the authority of one's elders, which implicitly regulates the interaction between the young and the old, is common practice (Liem, Martin, Nair, Bernardo, & Prasetya, 2009; Maulana et al., 2011). Consequently, the perceived distance between pupil and teacher is relatively high, suggesting that the school system tends to be hierarchical and monotonous. The combination of these unique cultural features may explain the importance of teachers as well as self-interest for pupils' learning activities in the Indonesian context.

Currently, there are some gaps in the knowledge base of teaching behaviour. First, research on teaching behaviour reviewed earlier is rather fragmented, focusing solely on one or some aspects of teaching behaviour. Second, the literature recognizes characteristics of effective teaching behaviour, but studies investigating characteristics of teaching behaviour that can be categorized into difficulty levels are scarce. It is important to determine the complexity level of behaviour to assist teachers develop in their zone of proximal development. Logically, we cannot expect teachers to perform more complex behaviour in their development without supporting them step by step from less complex behaviour. A measure covering more comprehensive aspects of teaching behaviour would be beneficial for facilitating the evaluation (and development) of teaching performance in a more comprehensive fashion. Such a robust measure is present, and was developed originally in the Dutch context (Maulana, Helms-Lorenz, & Van de Grift, 2014).

Our aim is to construct a robust diagnostic measure for evaluating teaching behaviour that can be put along a unidimensional continuum which varies in the level of difficulty for use in the Indonesian secondary education context.

Fig. 1. The conceptual model of the relationship between teaching behaviour, academic motivation, and outcomes. Adapted from Appleton, Christenson, and Furlong (2008).

Additionally, the predictive power of the teaching behaviour measure is the focus of this study as well. Particularly, the following research questions are of central interest:

1. To what extent is the measure of pupils' perceptions of teachers' teaching behaviour reliable and valid for use in secondary education in Indonesia?
2. How do pupil perceptions of teachers' teaching behaviour explain their perceived academic motivation? Is teaching behaviour important for both controlled and autonomous motivation of Indonesian pupils?

The conceptual model of the relationship between teaching behaviour and academic motivation tested in this study can be seen in Fig. 1.

4. Methods

4.1. Sample and procedure

The present study involved a total sample of 4000 pupils of 200 teachers/classes from 16 public schools in the West Java province, Indonesia. All schools were located in a provincial city area and are accessible for pupils from various socio-economic backgrounds. The present data is representative of typical public schools for middle-class pupils. The pupil sample included 1539 boys (39%) and 2459 girls (61%). Of the pupils, 50% were from junior high schools and the other half were from senior high schools. A total of 21% of the pupils were taught by their teachers for more than a school year, while the rest were taught for the first time during the school year.¹ All the teachers were certified to teach in secondary education. 38% of them were male and 62% were female. A total of 25% of the teachers taught math and natural science subjects, while the remaining 75% taught non-science subjects (e.g., social sciences, language).

Schools were recruited to participate in the survey on a voluntary basis. Letters were sent to principals of the schools to participate in the study. Upon official agreement to participate, questionnaires were distributed during the school year of 2013. In all schools, class sizes were large (ranging from 35 to 40 students). For the reason of representativeness and

¹ In Indonesia, annual rolling in teaching is possible. Therefore, teachers who teach the first grade during the first school year may be assigned to teach the second grade in the subsequent school year. Additionally, some teachers may teach two different grade levels during the same school year. That is why a relatively small number of pupils may have the same teacher for two consecutive school years.

parsimony, we selected 20 pupils per teacher/class randomly to participate in the survey.² Surveys were conducted by trained assistant researchers. In order to reduce response bias, teachers were not present in the class during the survey. It took approximately 20 min for pupils to complete the survey.

4.2. Measure

Prior to the data collection, the questionnaires measuring teachers' teaching behaviour and students' academic motivation for use in the Indonesian secondary school context were adapted. The English version of the questionnaires was used as the source language. The adaptation procedure was in accordance with the guidelines of the International Test Commission (Hambleton, 1994, 2001). A combination between the translation-backtranslation procedure and the committee approach was undertaken (Van de Vijver & Tanzer, 2004). With this combined approach, the quality of the translation is enhanced because linguistic, cultural, and psychological aspects are taken into account.

A group of people with different areas of expertise were involved in the adaptation process of the instruments. These included an educational researcher, together with two professional university interpreters majoring English as a Foreign Language (EFL) and one university professor in developmental psychology. The four experts were qualified and experienced in the source and target languages as well as in both cultures (The Netherlands and Indonesia). They were all knowledgeable about instrument development and adaptation. The first interpreter translated the questionnaires from English to Indonesian language. The university professor carefully checked the content and the appropriateness of the translation for use specifically in the context of secondary education. The second interpreter independently translated the questionnaires back into English. Finally, the committee discussed various kinds of errors found and resolved them subsequently. The committee applied the *successive* method of translation using the *application* approach (Van de Vijver & Tanzer, 2004). They agreed that the underlying constructs are appropriate in each cultural group and that more straightforward translations are sufficient for the majority of items to get an instrument that adequately measures the same construct in the target group.

4.3. Teaching behaviour

To measure pupils' perceptions of teachers' teaching behaviour, we used the My Teacher questionnaire previously developed by Maulana et al. (2014). The questionnaire was initially developed for use in the Dutch secondary school context and is available in Dutch and English versions. For the present study, the measure was adapted for the Indonesian context (see above the detail about the adaptation process). The measure consists of 24 items provided on a four-point Likert scale, ranging from 1 (Never) to 4 (Often). Consistent with the theoretical framework for teaching behaviour presented in the literature review earlier, the six observable domains are present in the measure (see Table 1).

4.4. Academic motivation

To assess pupils' perceived academic motivation, we used the Indonesian version of the Questionnaire on Motivational Dimension (Maulana et al., 2011; Maulana & Opdenakker, 2013). This questionnaire was based on the academic regulation scale of Ryan and Connell (1989). In line with self-determination theory, this measure taps pupils' perceived motivation in terms of *controlled* and *autonomous* motivation (Ryan & Deci, 2000). Conceptually, the scale of controlled motivation consists of the subscales of external (i.e., "I study this subject because my teacher asks me") and introjected (i.e., "I study this subject because I would feel ashamed if I don't do it") regulations, while the scale of autonomous motivation consists of the subscales of identified (i.e., "I study this subject because it's personally important to me") and intrinsic (i.e., "I study this subject because I find it interesting") motivation. The conceptual grouping of this motivational scales was tested statistically in terms of a reasonable simplex pattern by showing that each subscale should correlate more positively (or less negatively) with subscales closer to it and less positively (or more negatively) with subscales farther from it (Vansteenkiste et al., 2004).

The measure consists 16 items and four items per subscale, provided on a five-point Likert scale ranging from 1 (Completely disagree) to 5 (Completely agree). In agreement with the study of Maulana and Opdenakker (2013) in a smaller and less diverse sample of Indonesian secondary education, the reliability of this measure using the present data shows a satisfactory level (Cronbach's $\alpha_{\text{controlled motivation}} = 0.71$, Cronbach's $\alpha_{\text{autonomous motivation}} = 0.91$).

4.5. Analytic strategy

To answer the first research question, we used various classical test theory (CTT) and item response theory (IRT) approaches. For the CTT, we particularly used exploratory and confirmatory factor analyses (EFA and CFA). For the IRT, we specifically used the Rasch modelling. The combination of both approaches is a useful way to analyse the psychometric property of an instrument in a more comprehensive way. To answer the second research question, we performed multilevel modelling.

² A sample of about 15 pupils per class is generally sufficient to generate (aggregated) class perceptions of teachers' behaviour.

Table 1

Results of item response theory analyses of the teaching behaviour scale.

Domain	Item	Label	Item content My teacher...	Descriptive statistics and calibration					Item-level fit indices			
				<i>p</i>	<i>b</i>	<i>SE</i>	<i>a</i>	<i>se</i>	Q-index	Zq	S-X ²	<i>p</i>
Climate	1	V1	Treats me with respect	0.90	−0.98	0.18	0.73	0.21	0.29	1.26	68.35	0.00
Climate	2	V3	Answers my question	0.89	−0.88	0.18	1.04	0.23	0.23	0.76	26.62	0.04
Climate	3	V7	Ensures that I treat others with respect	0.88	−0.82	0.17	1.49	0.27	0.19	0.27	19.49	0.25
Instruction	4	V2	Prepares his/her lesson well	0.88	−0.79	0.17	2.23	0.37	0.12	−0.52	11.98	0.75
Differentiation	5	V8	Helps me if I don't understand or am unable to do something	0.87	−0.76	0.17	2.06	0.35	0.13	−0.38	14.24	0.59
Differentiation	6	V6	Helps me if I don't know or am unable to do something	0.87	−0.71	0.17	1.93	0.32	0.14	−0.29	13.82	0.62
Instruction	7	V11	Explains the purpose of the lesson	0.86	−0.65	0.17	1.93	0.32	0.15	−0.13	28.17	0.04
L. strategies	8	V9	Explains how I need to do things	0.84	−0.40	0.16	1.73	0.27	0.14	−0.09	19.71	0.29
Instruction	9	V5	Makes clear what I need to do for a test	0.83	−0.28	0.15	2.33	0.35	0.10	−0.67	10.37	0.89
L. strategies	10	V22	Explains how I should study something	0.82	−0.21	0.15	2.41	0.36	0.10	−0.66	12.92	0.75
Management	11	V16	Ensures that I pay attention	0.81	−0.08	0.15	1.60	0.24	0.15	0.26	20.01	0.27
Instruction	12	V10	Uses clear examples	0.80	−0.06	0.15	1.85	0.27	0.14	−0.05	13.03	0.68
Activation	13	V23	Motivates me	0.79	0.09	0.15	1.75	0.25	0.14	0.19	22.29	0.14
Activation	14	V12	Ensures that I do my best	0.78	0.13	0.14	2.11	0.30	0.11	−0.42	11.61	0.78
Instruction	15	V13	Explains everything clearly to me	0.77	0.17	0.14	2.14	0.30	0.11	−0.44	21.24	0.22
Differentiation	16	V17	Connects to what I know or am capable of	0.77	0.17	0.14	1.68	0.24	0.14	0.22	22.51	0.17
Management	17	V15	Applies clear rules	0.76	0.24	0.14	1.75	0.25	0.13	0.11	22.70	0.10
L. Strategies	18	V21	Stimulates my thinking	0.74	0.43	0.14	2.49	0.34	0.09	−0.76	21.17	0.18
Activation	19	V4	Involves me in the lesson	0.73	0.52	0.14	0.68	0.15	0.24	2.81	81.20	0.00
Differentiation	20	V20	Makes me feel self-confident with difficult tasks	0.72	0.58	0.14	2.26	0.30	0.09	−0.47	15.03	0.45
Differentiation	21	V18	Checks whether I have understood the subject matter	0.71	0.61	0.14	2.19	0.29	0.10	−0.46	19.58	0.19
Management	22	V19	Ensures that I use my time effectively	0.68	0.88	0.13	2.28	0.29	0.09	−0.51	36.09	0.00
Differentiation	23	V14	Pays attention to me	0.63	1.21	0.13	1.46	0.19	0.14	0.81	36.01	0.00
Differentiation	24	V24	Knows what I have difficulty with	0.57	1.61	0.13	2.77	0.35	0.07	−0.82	18.04	0.16

Note: *p* = proportion of correct response, *p* = significant level based on *S-X*² statistics, *b* = item difficulty level parameter from the Rasch model, *a* = slope parameter from the Birnbaum model (2-parameter logistic model), *S-X*² = item level diagnostic statistics, *se* = standard error.

4.6. Rasch Modelling

As a family of item response theory (IRT), the Rasch model (Rasch, 1960) has more stringent assumptions to be met compared to classical test theory (CTT). Edelen and Reeve (2007) argued that the usefulness of this of the IRT model depends on the extent to which these assumptions are met. For the Rasch model, there are three assumptions: (1) *unidimensionality*. This assumption requires that a set of items must measure one and the same latent skill in a homogenous way. (2) *Local stochastic independence*. This means that the response to one item may not influence the response to another, except for an influence that can be explained by the latent variable that is the measurement objective of the set of items. (3) *Parallelism of the item characteristic curves*. This means that each item in the Rasch scale should contribute uniquely to the scale. Items with extremely high or low discrimination power indicate a violation of this assumption (DeMars, 2010; Van de Grift, Van der Wal, & Torenbeek, 2011). When an instrument satisfies the three assumptions of the Rasch model, the instrument can reliably and validly be used for measuring the behavioural performance.

Prior to conducting the Rasch analysis, we conducted EFA and CFA to examine the unidimensionality and local stochastic independence assumptions. For this purpose, we divided the total sample into two parts: the first half was used for EFA, and the other half was used for CFA. All items were treated as dichotomous variables. To examine the unidimensionality from the EFA, we evaluated the scree plot, eigenvalues, and the magnitude of item loadings on the single-factor solution. Furthermore, a single-factor CFA with uncorrelated residuals model was examined. If a local dependency among pairs of items is evident, the single-factor model with uncorrelated residuals would not show an acceptable fit. For both EFA and CFA, goodness of model-data fit was evaluated based on four statistics indices: Chi-square test, the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA) (Schreiber, Nora, Stage, Barlow, & King, 2006).³

³ Fit indices criteria for model-data fit are: Chi-square test should not be significant (*p* < 0.05), the CFI and TLI value should be ≥ 0.95, and the RMSEA value should be between < 0.06 and 0.08 (Schreiber et al., 2006).

Fig. 2. Scree plot of the eigenvalues of the tetrachoric correlation matrix of the 24 items.

Furthermore, the Rasch model was conducted. Because the stability of estimation derived from Rasch model is sensitive to large sample size,⁴ we selected a sample of 400 cases out of the total sample randomly. With this sample, we fitted the Rasch model and calibrated the items subsequently. Firstly, the homogeneity of the set of items from the CTT perspective was examined in terms of the reliability, mean, variance of raw scores, and standard deviations. Next, the fit of the model at the item and the overall model fit was tested using the $S-X^2$ item-fit statistic (Orlando & Thissen, 2003). The item-fit indicator was also checked using the Q-index and the standardized Q-index (zq) (Rost & Von Davier, 1994). Thirdly, the local stochastic independency among pairs of items was tested using the $LD-X^2$ statistics proposed by Chen and Thissen (1997). Additionally, the slope of each item was also estimated using the two-parameter logistic model (Birnbaum, 1968).

The overall model-data fit was analysed further using the M_2 statistic and the RMSEA index (Maydeu-Olivares & Joe, 2005). Moreover, the parallelism of the item characteristic curve was checked using the Andersen (log-likelihood ratio) test (Andersen, 1977). Finally, the log-likelihood ratio test was also used to examine the unidimensionality dimension by comparing group differences in terms of teaching subject (science versus non-science) and student gender (boys versus girls). All analyses were done using various statistical programs including SPSS, WINMIRA (Von Davier, 1994), eRm from R (Mair, Hatzinger, & Maier, 2013), Mplus (Muthén & Muthén, 1998–2012), and IRTPRO (Cai, Thissen, & Du Toit, 2005–2013).

4.7. Multilevel modelling

In the present study, the data were naturally ordered in a hierarchical structure (pupils nested with classes). Therefore, multilevel modelling is a proper way to examine the relationship between teaching behaviour and academic motivation in a more precise way by taking into account the hierarchical structure of the data (Snijders & Bosker, 1999). We performed two-level multilevel analysis, with class/teacher at level 2 and pupils at level 1. Firstly, we estimated empty models for academic motivation (null model). Secondly, we included teaching behaviour as a predictor of academic motivation. Finally, several background variables were included to the model. Additionally, possible interaction effects between variables were checked as well. The modelling was done using a stepwise procedure. The estimation was done separately for controlled and autonomous motivation using a statistical program MLwiN (Rasbash, Charlton, Browne, Healy, & Cameron, 2005).

⁴ A sample as small as 100 is often adequate for estimation of the Rasch model (Edelen & Reeve, 2007).

Table 2
Standardized $LD-X^2$ diagnostic statistics for the teaching behaviour scale.

Label	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	–																						
2	–0.6	–																					
3	–0.1	–0.3	–																				
4	3.1	6.4	6.8	–																			
5	1.7	0.2	–0.7	3.5	–																		
6	–0.6	–0.7	0.3	5.5	3.4	–																	
7	2.7	–0.1	3.5	16.7	0.7	–0.7	–																
8	0.4	–0.7	–0.0	8.6	0.0	6.4	–0.4	–															
9	6.8	–0.2	0.9	–0.6	–0.6	–0.3	3.3	–0.7	–														
10	6.1	0.2	1.6	9.4	0.3	0.3	–0.3	0.5	0.4	–													
11	5.6	0.6	3.8	8.1	0.3	2.1	3.2	–0.7	1.2	0.1	–												
12	4.1	–0.3	4.2	7.3	–0.6	–0.5	–0.6	–0.1	–0.7	0.7	–0.3	–											
13	7.3	–0.5	2.4	8.2	0.8	–0.4	0.1	–0.4	–0.6	5.0	–0.6	–0.4	–										
14	4.5	1.7	6.9	0.3	–0.4	1.5	3.1	7.2	0.3	6.2	2.5	–0.4	1.5	–									
15	10	–0.2	8.5	11.7	–0.4	5.9	–0.6	4.9	–0.6	5.6	0.2	–0.1	1.1	0.5	–								
16	7.4	1.5	12.8	13.5	2.0	5.8	–0.0	6.3	0.3	2.9	3.0	–0.7	–0.4	–0.5	3.2	–							
17	7.3	–0.1	1.5	15.9	0.2	–0.0	2.2	–0.3	–0.1	–0.5	0.9	–0.2	–0.3	1.5	–0.2	–0.4	–						
18	1.8	–0.2	3.2	13.1	–0.0	–0.2	0.7	0.1	0.5	–0.5	0.3	–0.1	–0.6	–0.5	–0.6	–0.6	–0.1	–					
19	1.0	–0.5	3.4	14.6	0.3	–0.1	–0.5	0.3	0.0	0.7	0.4	4.6	–0.3	–0.1	–0.5	–0.5	–0.5	1.9	–				
20	4.8	–0.6	2.0	12.1	–0.6	–0.3	3.9	–0.3	–0.1	–0.2	–0.5	–0.5	–0.5	1.5	0.4	–0.5	0.4	–0.4	1.2	–			
21	2.4	–0.6	1.8	8.3	–0.6	–0.6	3.6	–0.6	–0.6	0.2	0.2	–0.5	–0.2	–0.4	–0.6	–0.5	–0.6	–0.2	–0.2	3.6	–		
22	2.2	–0.6	4.4	9.9	0.3	–0.3	1.1	–0.5	–0.6	–0.6	–0.6	–0.3	–0.4	0.1	–0.6	–0.7	2.8	–0.4	0.5	–0.6	6.3	–	
23	3.7	–0.6	2.7	16.9	3.0	–0.7	0.9	–0.6	0.9	3.3	–0.0	–0.6	0.7	2.1	1.5	2.1	3.7	–0.5	0.0	0.2	–0.3	–0.4	–
24	10	–0.4	5.2	7.0	3.9	–0.5	1.9	–0.5	–0.4	–0.3	–0.3	0.0	–0.1	–0.3	1.0	–0.2	–0.5	5.2	8.6	10	3.5	0.2	1.7

Note. Values range <5 indicates likely no local dependency, 5–10 is in the grey area; may indicate either local dependency or the sparseness of the frequency table, >10 indicates likely local dependency.

Fig. 3. The item characteristic curves of the 24 items.

5. Results

5.1. The psychometric quality of teaching behaviour measure

Results of the single-factor solution of EFA shows a reasonable good fit ($\chi^2_{(1996)} = 757.49$, $p = 0.00$, CFI = 0.97, TLI = 0.96, RMSEA = 0.03). The scree plot of eigenvalues is strongly suggestive of a single factor as well, with the first value (12.21) significantly larger than the others (1.40, 0.91, 0.82, etc. see Fig. 2). The values of factor loadings are positive for all items, ranging from 0.46 to 0.81. All items have factor loadings above 0.70, except for item 1 (“My teacher treats me with respect (0.58)”) and item 19 (“My teacher involves me in the lesson (0.46)”). Results of EFA suggests that the teaching behaviour construct is sufficiently unidimensional. Results of CFA confirms that the one-factor solution model is a good representation of the data as well ($\chi^2_{(1933)} = 1282.019$, $p = 0.00$, CFI = 0.98, TLI = 0.97, RMSEA = 0.04). Based on results of EFA and CFA, we conclude that the unidimensionality of teaching behaviour construct holds in the present sample.

Results of classical test analysis indicate that the 24 items of teaching behaviour measure is sufficiently homogeneous (Reliability = 0.80, $\text{Mean}_{(\text{raw scores})} = 18.96$, $\text{SD} = 5.05$). Results of the Rasch modelling can be seen in Table 1. Of all 24 items, $S-X^2$ item level diagnostic statistics show that four items show a significant value at $p < 0.01$. Those items are: item 1 (“My teacher treats me with respect”), item 19 (“My teacher involves me in the lesson”), item 22 (“My teacher ensures that I use my time effectively”), and item 23 (“My teacher pays attention to me”). Furthermore, the Q-index ranges between 0.07 (item 24) and 0.29 (item 1).⁵ Based on this index, there is no significant deviation for the expected characteristic of the Rasch model. The corresponding Zq-index ranges between −0.82 (item 24) and 2.81 (item 19). Item 1 (“My teacher treats me with respect”) and item 19 (“My teacher involves me in the lesson”) have rather high positive values, suggesting that these two items appear to have lower discrimination power than assumed by Rasch model (under-fit).

Furthermore, results of $LD-X^2$ suggest that item 2 (“My teacher answers my question”) seems to be locally dependent with item 11 (“My teacher ensures that I pay attention”, see Table 2).⁶ Item 19 (“My teacher involves me in the lesson”) seems to be locally dependent with items 3, 11, 13, 16, 17, 20, 21, and 22. This indicates that item 2 and item 19 appear to violate the local stochastic independency assumption. Results from the Birnbaum model (2-parameter logistic model) show that the 24 items have a slope between 0.68 (item 19, “My teacher involves me in the lesson”) and 2.77 (item 24, “My teacher knows what I have difficulty with”). None of the items has a high slope value (e.g., >4) to be considered excessively dependence (Edelen & Reeve, 2007). However, the slope values of item 1 and item 19 seem to be rather low compared to the rest of the items.

Based on inspection of the overall model-data fit, we found that the M_2 statistics indicates a lack of fit ($M_{2(275)} = 676.03$, $p = 0.00$). The associated RMSEA value is 0.06, suggesting that the lack of fit is possibly due to a minor amount of “model error”. Therefore, the overall model-data fit is deemed acceptable. Additionally, the result of Andersen test comparing teachers with high and low teaching behaviour performances is not significant ($\chi^2_{(23)} = 16.81$, $p = 0.78$). This suggests that the item characteristic curves of the 24 items are reasonably parallel (see Fig. 3). As a final check of unidimensionality assumption, results of the Andersen test show based on comparisons between teaching subject and student gender are not significant at $p < 0.01$ ($\chi^2_{(23)} = 33.21$, $p = 0.08$; $\chi^2_{(23)} = 38.63$, $p = 0.03$). From these results, we ascertain that the

⁵ The Q-index can vary between 0 (perfect fit, i.e., a Guttman pattern) and 1 (perfect misfit), with 0.5 indicating no relationship between the individual parameter and the reaction to the item. High positive values of the zq index indicate that the item discrimination is lower than assumed by the Rasch model (under-fit), while high negative values indicate higher discrimination than assumed (over-fit).

⁶ $LD-X^2 > 10$ indicates likely locally dependence among pairs of items, while the value between 5 and 10 lies in the grey area, suggesting that it may either indicate local dependencies or maybe a result of sparseness in the underlying table of frequencies. $LD-X^2 < 5$ indicates likely no local dependencies (Scientific Software International, 2011).

Table 3

Results of multilevel analysis to examine the link between teaching behaviour and academic motivation; parameter estimates.

Variable	Academic motivation							
	Controlled motivation				Autonomous motivation			
	Model 0		Model 1		Model 0		Model 1	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
<i>Fixed effects</i>								
Intercept	3.297***	0.018	2.455***	0.090	3.842***	0.019	2.377***	0.096
Teaching behaviour			0.301***	0.024			0.471***	0.026
Teaching subject			−0.056	0.040			−0.037	0.043
Teacher gender			−0.031	0.035			−0.013	0.037
Grade level			0.009	0.035			−0.007	0.037
Student gender			−0.022	0.021			0.071**	0.023
Duration taught			−0.006	0.035			0.030	0.037
<i>Random effects</i>								
Class/teacher level	0.047	0.006	0.040	0.006	0.049	0.007	0.043	0.006
Student level	0.353	0.008	0.343	0.008	0.437	0.010	0.404	0.009
Deviance	7415.692		7153.046		8246.479		7788.447	

* $p < 0.05$.** $p < 0.01$.*** $p < 0.001$.

unidimensionality of teaching construct is not violated by teaching subject and student gender. Finally, the measure of teaching behaviour has a good predictive validity. Teaching behaviour is significantly and positively correlated with controlled motivation ($r = 0.31$, $p < 0.00$) and with autonomous motivation ($r = 0.35$, $p < 0.00$).

5.2. Teaching behaviour and academic motivation

Based on results of multilevel analyses for controlled motivation, we found that 12% of the variance is attributed to class/teacher level and 88% of the variance is attributed to pupil level. For autonomous motivation, we found that 10% of the variance is accounted for class/teacher level and 90% is accounted for pupil level (see Table 3, Model 0). This means that between class as well as between pupil differences in controlled and autonomous motivation are visible. Compared to differences between classes, differences between pupils in academic motivation are estimated to be far larger.

We found that pupils' perceptions of effective teaching behaviour are significantly and positively related to their perceived academic motivation, even after controlling for background variables (see Table 3, Model 1). Results suggest that effective teaching behaviour can significantly predict controlled ($\beta = 0.30$, $p < 0.00$) and autonomous motivation ($\beta = 0.47$, $p < 0.00$). Effective teaching behaviour explains about 8% of the variance in autonomous motivation and about 4% of the variance in controlled motivation. This suggests that the magnitude of the relationship between effective teaching behaviour and autonomous motivation is stronger than that of the relationship between effective teaching behaviour and controlled motivation. Additional analyses reveal that there is no interaction effect between the effects of effective teaching behaviour and academic motivation when the background variables are taken into account.

6. Conclusions and discussion

Teaching behaviour has been recognized as a catalyst for pupils' learning and outcomes. However, the body of knowledge is still limited in at least three particular areas. First, existing research and measure of teaching behaviour are rather fragmented, paying attention merely to solo, or very limited aspects of teaching behaviour. Furthermore, studies investigating characteristics of teaching behaviour in terms of the level of difficulty is scarce. Additionally, studies on the effect of teaching behaviour on pupils' academic motivation are predominantly originated from the Western context. In the non-Western context (i.e., Indonesia) this type of study is still underrepresented. The present study addresses these gaps by evaluating the usefulness of an instrument tapping teaching behaviour originated from the Western context, as well as investigating the effect of teaching behaviour on pupils' controlled and autonomous motivation from Indonesian pupils' point of view.

We found that the measure of pupils' perceptions of teachers' teaching behaviour reliable and valid for use in secondary education in Indonesia. Result of classical test analysis showed that the measure is highly reliable. More importantly, results of Rasch analyses revealed that the measure satisfies the stringent assumptions required by the Rasch model reasonably well. A product of an evaluation instrument that satisfies the Rasch model is a robust tool for evaluating the teaching performance with high accuracy (Edelen & Reeve, 2007). This finding is consistent with the research of Maulana et al. (2014) which showed that the original Dutch version of this teaching behaviour measure meets the assumptions of the Rasch model.

This suggests that the Indonesian version of the teaching behaviour measure can be used as a robust instrument that can function as a diagnostic tool for improving teaching behaviour targeting at the zone of proximal development in the country.

With the Rasch score, it is possible to determine the level of teachers' teaching behaviour at certain moments. Information about the level of teaching behaviour at certain moments can be used as an indicator for determining teachers' zone of proximal development. This approach was done initially in the effort to support professional development of teachers in the Netherlands (Van de Grift, Maulana, & Helms-Lorenz, 2014). With the Rasch modelling approach, they were able to determine the level of teachers' teaching behaviour based on differences in teaching experiences. They found that experienced teachers are better than inexperienced teachers. While experienced teachers already master more complex teaching skills (i.e., activating learning and teaching strategy), inexperienced teachers are still struggling with less complex teaching skills (i.e., stimulating learning climate and classroom management; Van de Grift et al., 2014a). Supporting teachers to improve their teaching behaviour from the zone of proximal development is a logical way to undertake because school leaders, school evaluators, mentors, and coaches can help teachers through progressive steps. Having said that, teachers need to be supported step by step from skills that they already acquire and move through more complex skills subsequently. This idea is in line with the importance of scaffolding as a means to support (teachers') learning (Hattie & Clinton, 2008; Houtveen & Van de Grift, 2007). Rationally, it might be misleading to expect teachers acquire more complex teaching skills before they attain less complex ones.

Furthermore, we found that pupil perceptions of teachers' teaching behaviour explained their perceived academic motivation significantly and positively. Results indicated that a higher level of teaching behaviour corresponds to a higher level of academic motivation. This finding is consistent with past research, predominantly from the Western context, showing the importance of teaching behaviour for pupils' learning motivation (e.g., Kunter et al., 2007; Kyriakides et al., 2009; Pavey et al., 2011; Stroet et al., 2013; Wentzel, 1999). Importantly, we found that teaching behaviour has positive and significant effects on both controlled as well as autonomous motivation. This finding suggests that teachers seem to be powerful to influence the motivation in learning of Indonesian pupils in a controlled and autonomous way. Our finding is in line with the argument of Asian scholars stressing the importance of significant others (i.e., parents) as a reason for pupils to learn (Tseng, 2004; Zhu & Leung, 2010), which reflects a collectivist culture associated with respect for the authority of one's elders (Liem et al., 2009; Maulana et al., 2011). Indonesian culture highly values obedience to an authority and respected figure like teachers. The fact that teachers have a powerful role in influencing their pupils to learn for external reasons might represent a unique feature of Indonesian classroom practices (and classrooms of the neighbouring countries sharing similar cultural values), which can be contrasted with Western classroom practices. Indeed, past research from the Dutch context showed that teachers did not have a powerful role to influence pupils' controlled motivation (Opdenakker et al., 2012).

However, we found that the effect of teaching behaviour on controlled motivation is, although significant, not as strong as that on autonomous motivation. This finding signifies the importance of teaching behaviour for pupils' inner interest in learning, which is highly valued by both Western and Eastern contexts. The initial idea of the importance of social context (i.e., teachers) for pupils' interest in learning is originated from self-determination theory (Ryan & Deci, 2000), a theory which was born and becoming more popular in the Western context. Self-determination theory posits that teaching behaviour should be directed towards the facilitation of learning for inner interest (self-determined) in order to sustain pupils' learning development productively. The fact that teaching behaviour is more strongly related to Indonesian pupils' autonomous motivation provides a support that self-determination theory is a universal theory: Teachers are universally important for stimulating pupils' self-interest in learning, regardless of the cultural context. Another support for this argument can also be found in the research of Maulana and Opdenakker (2013) and Maulana et al. (2013) who showed that teacher involvement, friendliness, and interpersonal dominance are beneficial for pupils' self-interest in learning in Indonesia. Nevertheless, more studies in the Asian context are needed to validate our conclusion further.

In addition to the scientific importance mentioned above, the present study has an important implication for practice as well. Finding a robust instrument to evaluate teaching performance can considerably be tricky and costly. However, the quality of an evaluation instrument is a crucial factor determining the effectiveness of teaching evaluation aimed for supporting teachers to develop their professional teaching skills continuously through their teaching career. The instrument presented in this paper meets the strict requirements of a robust measure. The robustness was shown from the Rasch modelling and predictive value perspective. Practically, this instrument can be used widely as a guideline to support teachers' teaching behaviour targeting at their zone of proximal development. Schools are suggested to use the instrument for determining the level of their teachers' teaching performance. In cooperation with school leaders and educational developers, teachers can be supported to develop their professional teaching behaviour stage-wisely, focusing first on less complex behaviours and moving forward to more complex ones. Moreover, policy makers can use this instrument as an evaluation measure to assess the quality of teachers regionally and nationally. Information about the level of teachers' teaching performance nationally can be a useful information for policy makers so appropriate attention and intervention can be devoted to the specific target group. A continuous development promotes the enhancement of educational improvement.

The benefit of pupils' autonomous motivation is not only for pupils themselves (i.e., high achievement), but also for teachers. A research in Korea indicated that a lack of interest in learning has been shown to be one of the most detrimental factors for teachers (Kim, Kim, & Zhang, 2014). As the present study showed that high levels of academic motivation corresponds to high levels of teaching behaviour, we argue that teachers should meet a certain level of teaching performance before they enter the teaching profession. Setting the standard and minimum level should be done during the training in the

teacher institute and improving teaching behaviour should be a priority in the teacher continuous professionalism agenda. At a certain point, when teachers are able to influence pupils' motivation, teachers will likely to experience the joy of teaching. Additionally, efforts to enhance the quality of education will be optimized through international cooperation.⁷ It is certainly beneficial to learn from each other's success of improving the quality of teaching practices. Hence, it is our hope that this paper would stimulate global discussion and cooperation regarding the improvement of teachers' teaching quality among researchers in the Asia-Pacific region.

Regardless of the strengths, the present study also subjects to limitations. The study took place in Indonesia with samples coming solely from West Java. Therefore, interpretation of findings regarding Western and non-Western differences should be handled with caution until research involving wider non-Western samples is available. Nevertheless, the current study involved a relatively large sample size allowing us to draw conclusions about the usefulness of the measure of teaching behaviour. Furthermore, the present study assumes a one-directional effect of teaching behaviour on motivation. In practice, the direction of the effect can be the opposite: Motivation affects teaching behaviour as well. Although the present study signifies the importance of teaching behaviour for pupils' motivation, the body of knowledge can benefit from future studies addressing the reciprocal effect between teaching behaviour and motivation. Additionally, interpretations of our findings are limited in terms of perceptions of pupils. (Natural) classroom observations on classroom practices would be worth doing to validate and enrich our findings regarding "real" classroom practices and effects on pupils' learning. Indeed, enhancing the quality of education requires multiple approaches involving multiple elements working together cooperatively.

References

- Andersen, E. B. (1977). Sufficient statistics and latent trait models. *Psychometrika*, 42, 69–81.
- Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, 45, 369–386.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397–479). Reading, MA: Addison-Wesley.
- Cai, L., Thissen, D., & Du Toit, S. (2005–2013). *IRTPRO (Version 2.1) [Computer software]*. Lincolnwood, IL: Scientific Software International.
- Carnine, D. W., Dixon, R. C., & Silbert, J. (1998). Effective strategies for teaching Mathematics. In E. J. Kameenui & D. W. Carnine (Eds.), *Effective teaching strategies that accommodate diverse learners*. Upper Saddle River, NJ: Prentice Hall.
- Chen, W.-H., & Thissen, D. (1997). Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*, 22, 265–289.
- Chirkov, V. I., & Ryan, R. M. (2001). Parent and teacher autonomy support in Russian and U.S. Adolescents: Common effects on well-being and academic motivation. *Journal of Cross-Cultural Psychology*, 32, 618–635.
- Chye, S. Y., Liau, A. K., & Liu, W. C. (2012). Student teachers' motivation and perceptions of E-portfolio in the context of problem-based learning. *Asia-Pacific Education Researcher*. <http://dx.doi.org/10.1007/s40299-012-0022-4>
- Cornelius-White, J. (2007). Learner-centred teacher-student relationships are effective: A meta-analysis. *Review of Educational Research*, 77(1), 113–143.
- Creemers, B. P. M. (1994). *The effective classroom*. London: Cassell.
- Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. London: Routledge.
- Danielson, C. (2013). *The framework for teaching: Evaluation instrument*. Princeton: The Danielson Group.
- De Jong, R., & Westerhof, K. (2001). The quality of student ratings of teacher behaviour. *Learning Environments Research*, 4, 51–85.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behaviour*. New York: Plenum.
- DeMars, C. (2010). *Items response theory: Understanding statistics measurement*. New York, NY: Oxford University Press.
- den Brok, P., Brekelmans, M., & Wubbels, T. (2004). Interpersonal teacher behaviour and student outcomes. *School Effectiveness and School Improvement*, 15, 407–442.
- den Brok, P., Fisher, D., & Scott, R. (2005). The importance of teacher interpersonal behavior on students' subject-specific motivation. *Journal of Classroom Interaction*, 40, 20–33.
- Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modelling to questionnaire development, evaluation, and refinement. *Quality of Life Research*, 16(1S), 5–18.
- Fraser, B. (1991). Two decades of classroom environment research. In H. J. Walberg (Ed.), *Educational environments: Evaluations, antecedents and consequences* (pp. 3–27). Elmsford, NY: Pergamon.
- Fraser, B. J., & Fisher, D. L. (1982). Predicting students' outcomes from their perceptions of classroom psychosocial environment. *American Educational Research Journal*, 19, 498–518.
- Gottfried, A. W., Gottfried, A. E., Cook, C. R., & Morris, P. E. (2005). Educational characteristics of adolescents with gifted academic intrinsic motivation: A longitudinal investigation from school entry through early adulthood. *Gifted Child Quarterly*, 49(2), 172–186.
- Good, T. L., & Brophy, J. (1989). Teaching the lesson. In R. E. Slavin & 1 (Eds.), *School and classroom organization* (pp. 25–68). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment (Bulletin of the International Test Commission)*, 10, 229–244.
- Hambleton, R. K. (2001). The next generation of the ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment*, 17, 164–172.
- Hampton, S. E., & Reiser, R. A. (2004). Effects of a theory-based feedback and consultation process on instruction and learning in college classrooms. *Research in Higher Education*, 45, 497–527.
- Hattie, J. A. C. (2012). *Visible learning for teachers: Maximizing impact on learning*. Abingdon and New York: Routledge.
- Hattie, J. A. C., & Clinton, J. (2008). Identifying accomplished teachers: A validation study. In L. Ingvarson & J. A. C. Hattie (Eds.), *Assessing teachers for professional certification: The first decade of the National Board for Professional Teaching Standards* (pp. 313–344). Oxford, UK: Elsevier.
- Helms-Lorenz, M., Slof, B., & Van de Grift, W. (2013). First year effects of induction arrangements on beginning teachers' psychological processes. *European Journal of Psychology of Education*, 28, 1265–1287.
- Hofstede, G. (1991). *Cultures and organizations: Software of the mind: Intercultural cooperation and its important for survival*. New York: McGraw-Hill.
- Houtveen, A. A. M., Booij, N., De Jong, R., & Van de Grift, W. J. C. M. (1999). Adaptive instruction and pupil achievement. *School Effectiveness and School Improvement*, 10(2), 172–192.

⁷ To readers who are interested in participating in an international comparison of teaching quality using the instrument presented in this paper, please feel free to contact the authors.

- Houtveen, A. A. M., & Van de Grift, W. J. C. M. (2007). Effects of meta cognitive strategy instruction and instruction time on reading comprehension. *School Effectiveness and School Improvement*, 18(2), 173–190.
- Kame'enui, E. J., & Carnine, D. W. (Eds.). (1998). *Effective teaching strategies that accommodate diverse learners*. Upper Saddle River, NJ: Prentice Hall.
- Kim, T.-Y., Kim, Y.-K., & Zhang, Q.-M. (2014). Differences in demotivation between Chinese and Korean English teachers: A mixed-methods study. *Asia Pacific Education Researcher*, 23, 299–310.
- Kindsvatter, R., Wilen, W., & Ishler, M. (1988). *Dynamics of effective teaching*. New York: Longman.
- Koeze, P. A. (2007). *Differentiated instruction: The effect on student achievement in an elementary school* (Unpublished doctoral dissertation) Ypsilanti, MI: Eastern Michigan University.
- Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht*. [Multiple goals in mathematics education]. Münster, Germany: Waxmann.
- Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning & Instruction*, 17(5), 494–509.
- Kyriakides, L., Creemers, B. P. M., & Antoniou, P. (2009). Teacher behavior and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education*, 25, 12–23.
- Lang, J., & Kersting, M. (2007). Regular feedback from student ratings of instruction: Do college teachers improve their ratings in the long run? *Instructional Science*, 35(3), 187–205.
- Leung, F. K. S. (2001). In search for East Asian identity in mathematics education. *Educational Studies in Mathematics*, 47, 35–51.
- Liem, G. A. D., Martin, A. J., Nair, E., Bernardo, A. B. I., & Prasetya, P. H. (2009). Cultural factors relevant to secondary school students in Australia, Singapore, Philippines and Indonesia: Relative differences and congruencies. *Australian Journal of Guidance and Counseling*, 19, 161–178.
- Lundberg, L., & Linnakylä, P. (1992). *Teaching reading around the world*. The Hague: IEA.
- Mair, P., Hatzinger, R., & Maier, M. J. (2013). *Extended Rasch modelling*. Retrieved from <http://cran.r-project.org/web/packages/eRm/vignettes/eRm.pdf>
- Marzano, R. J. (2003). *What works in schools: Translating research into action*. Alexandria, VA: ASCD.
- Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2014). Development and evaluation of a questionnaire measuring pre-service teachers' teaching behavior: A Rasch modelling approach. *School Effectiveness and School Improvement*. <http://dx.doi.org/10.1080/09243453.2014.939198> Advance online publication
- Maulana, R., & Opendakker, M.-C. (2013). Teachers' interpersonal involvement as a predictor of students' academic motivation among Indonesian secondary school students: A multilevel growth curve analysis. *Asia Pacific Education Researcher*. <http://dx.doi.org/10.1007/s40299-013-0132-7>
- Maulana, R., Opendakker, M.-C., & Bosker, R. (2013). Teacher-student interpersonal relationships do change and affect academic motivation: A multilevel growth curve modelling. *British Journal of Educational Psychology*, 84, 459–482.
- Maulana, R., Opendakker, M.-C., den Brok, P., & Bosker, R. (2011). Teacher-student interpersonal relationships in Indonesia: Profiles and importance to student motivation. *Asia Pacific Journal of Education*, 31(1), 33–49. <http://dx.doi.org/10.1080/02188791.2011.544061>
- Maulana, R., Opendakker, M.-C., Stroet, K., & Bosker, R. (2012). Observed lesson structure during the first year of secondary education: Exploration of change and link with academic engagement. *Teaching and Teacher Education*, 28, 835–850.
- Maydeu-Olivares, A., & Joe, H. (2005). Limited-and-full-information estimation and goodness of fit testing in 2n contingency tables: A unified framework. *Journal of the American Statistical Association*, 100, 1009–1020.
- Meeuwisse, M., Severiens, S. E., & Born, M. P. (2010). Learning environment, interaction, sense of belonging and study success in ethnically diverse student groups. *Research in Higher Education*, 51(6), 528–545.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School matters: The junior years*. Somerset: Open Books.
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education*, 7, 133–144.
- Nunes, T., & Bryant, P. (1996). *Children doing mathematics*. Oxford: Blackwell.
- Olsen, S. F., Yang, C., Hart, C. H., Robinson, C. C., Wu, P., Nelson, D. A., et al. (2002). Maternal psychological control and preschool children's behavioral outcomes in China, Russia, and the United States. In B. K. Barber (Ed.), *Intrusive parenting: How psychological control affects children and adolescents* (pp. 235–262). Washington, DC: American Psychological Association.
- Opendakker, M.-C., Maulana, R., & den Brok, P. (2012). Teacher-student interpersonal relationships and academic motivation within one school year: Developmental changes and linkage. *School Effectiveness and School Improvement*, 23, 95–119. <http://dx.doi.org/10.1080/09243453.2011.619198>
- Opendakker, M.-C., & Minnaert, A. (2011). Relationship between learning environment characteristics and academic engagement. *Psychological Reports*, 109(1), 259–284. <http://dx.doi.org/10.2466/09.10.11.PRO.109.4.259-284>
- Orlando, M., & Thissen, D. (2003). Further investigation of the performance of S-X²: An item fit index for use with dichotomous item response theory models. *Applied Psychological, Measurement*, 27, 289–298.
- Pavey, L., Greitemeyer, T., & Sparks, P. (2011). Highlighting relatedness promotes prosocial motives and behavior. *Personality and Social Psychology Bulletin*, 37, 905–917.
- Pearson, P. D., & Fielding, L. (1991). Comprehension instruction. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 815–860). White Plains, NY: Longman.
- Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, 8(30), 317–344.
- Peetsma, T., & Van der Veen, I. (2011). Relations between the development of future time perspectives in three life domains, investment in learning, and academic achievement. *Learning and Instruction*, 21, 481–494.
- Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38, 109–119.
- Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge. Attempting to construct explanatory answers facilitates learning. *Educational Psychologist*, 27(1), 91–109.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Denmark's Paedagogiske Institut.
- Rasbash, J., Charlton, C., Browne, W. J., Healy, M., & Cameron, B. (2005). *MLwiN (Version 2.0) [Computer software]*. Bristol, UK: University of Bristol.
- Reeve, J. (2002). Self-determination theory applied to educational settings. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 183–203). Rochester, NY: University of Rochester Press.
- Reis, S. M., McCoach, D. B., Little, C. A., Muller, L. M., & Kaniskan, R. B. (2011). The effects of differentiated instruction and enrichment pedagogy and reading achievement in five elementary schools. *American Educational Research Journal*, 48, 462–501.
- Rosenshine, B. (1980). How time is spent in elementary classrooms. In C. Denham & A. Lieberman (Eds.), *Time to learn* (pp. 107–126). Washington, DC: National Institute of Education.
- Rosenshine, B., & Berliner, D. (1978). Academic engaged time. *British Journal of Teacher Education*, 4, 3–16.
- Rosenshine, B. V., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 376–390). New York: Macmillan.
- Rosenshine, B. V., & Meister, C. (1997). Cognitive strategy instruction in reading. In S. Stahl & D. A. Hayes (Eds.), *Instructional models in reading* (pp. 85–107). Mahwah, NJ: Erlbaum.
- Rost, J., & Von Davier, M. (1994). A conditional item fit index for Rasch models. *Applied Psychological, Measurement*, 18, 171–182.
- Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57, 749–761.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68–78.
- Scantlebury, K., Boone, W., Kahle, J. B., & Fraser, B. J. (2001). Design, validation, and use of an evaluation instrument for monitoring systemic reform. *Journal of Research in Science Teaching*, 38, 646–662.
- Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.

- Schick, H., & Phillipson, S. N. (2009). Learning motivation and performance excellence in adolescents with high intellectual potential: What really matters? *High Ability Studies*, 20(1), 15–37.
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modelling and confirmatory factor analysis results: A review. *Journal of Educational Research*, 99, 323–337.
- Scientific Software International (2011). *IRTPRO users' guide [Computer software]*. Lincolnwood, IL: Scientific Software International.
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77, 454–499.
- Sijstra, J. (1997). *Balans van het taalonderwijs aan het einde van de basisschool: Uitkomsten van de tweede taalpeiling einde basisonderwijs*. [Achievements in mother tongue at the end of primary education]. Arnhem: CITO.
- Slavin, R. (1996). *Education for all*. Exton, PA: Swets & Zeitlinger Publishers.
- Smith, T. W., Baker, W. K., Hattie, J. A. C., & Bond, L. (2008). A validity study of the certification system of the National Board for Professional Teaching Standards. In L. Ingvarson & J. A. C. Hattie (Eds.), *Assessing teachers for professional certification: The first decade of the National Board for Professional Teaching Standards* (pp. 345–378). Oxford, UK: Elsevier.
- Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advance multilevel modeling*. London, England: Sage.
- Sternberg, R. J. (2001). Giftedness as developing expertise: A theory of the interface between high abilities and achieved excellence. *High Ability Studies*, 12(2), 159–179.
- Stroet, K., Opendakker, M.-C., & Minnaert, A. (2013). Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educational Research Review*, 9, 65–87.
- Teodorovic, J. (2011). Classroom and school factors related to student achievement: What works for students? *School Effectiveness and School Improvement*, 22, 215–236.
- Thomson, A., & Gaudreau, P. (2008). From optimism and pessimism to coping: The medicating role of academic motivation. *International Journal of Stress Management*, 15, 269–288.
- Tomlinson, C. (2003). *Fulfilling the promise of the differentiated classroom: Tools and strategies for responsive teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Tomlinson, C. (2005). Grading and differentiation: Paradox or good practice? *Theory into Practice*, 44, 262–269.
- Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview.
- Tseng, V. (2004). Family interdependence and academic adjustment in college: Youth from immigrant and U.S.-born families. *Child Development*, 75, 966–983.
- Van de Grift, W. J. C. M. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49, 127–152.
- Van de Grift, W. J. C. M. (2013). Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*. <http://dx.doi.org/10.1080/09243453.2013.794845> Advance online publication
- Van de Grift, W. J. C., & Helms-Lorenz, M. (2012). *Vaardigheid en ervaring van leraren*. [Teaching skills and teacher experience]. (submitted for publication).
- Van de Grift, W., Helms-Lorenz, M., & Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Studies in Educational Evaluation*, 43, 150–159.
- Van de Grift, W. J. C. M., Maulana, R., & Helms-Lorenz, M. (2014, January). *Beginning teachers' wellbeing, teaching skills and job leaving: Longitudinal development and effects of an experiment*. Paper presented at the international conference of school effectiveness and improvement, Yogyakarta, Indonesia.
- Van de Grift, W. J. C. M., Van der Wal, M., & Torenbeek, M. (2011). Ontwikkeling in de pedagogisch didactische vaardigheid van leraren in het basisonderwijs [Development of teachers' teaching skills in primary education]. *Pedagogische Studiën*, 88, 416–432.
- Van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 54, 119–135.
- Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., & Deci, E. L. (2004). Motivating learning, performance, and persistence: The synergistic role of intrinsic goals and autonomy-support. *Journal of Personality and Social Psychology*, 87, 246–260.
- Von Davier, M. (1994). *WINMIRA 2001 [Computer software]*. Kiel, Germany: IPN.
- Wang, M. C., Reynolds, M. C., & Walberg, H. J. (1995). Serving students at the margins. *Educational Leadership*, 52(4), 12–17.
- Wentzel, K. R. (1999). Social-motivational processes and interpersonal relationships: Implications for understanding students' academic success. *Journal of Educational Psychology*, 91, 76–97.
- Willms, J. D., & Somers, M.-A. (2001). Family, classroom, and school effects on children's educational outcomes in Latin America. *School Effectiveness and School Improvement*, 12, 409–445.
- Wubbels, T., & Brekelmans, M. (2005). Two decades of research on teacher-student relationships in class. *International Journal of Educational Research*, 43, 6–24.
- Yair, G. (2000). Educational battlefields in America: The tug-of-war over students' engagement with instruction. *Sociology of Education*, 73(4), 247–269.
- Zhu, Y., & Leung, F. K. S. (2010). Motivation and achievement: Is there an East Asian model? *International Journal of Science and Mathematics Education*, 9, 1189–1212.