Teachers' interpersonal relationships and instructional expertise: how are they related?

Lans, R.M. van der; Cremers, J.; Klugkist, I.; Zwart, R.

Studies in Educational Evaluation 66 (2020) 100902

Contents lists available at ScienceDirect

Studies in Educational Evaluation

journal homepage: www.elsevier.com/locate/stueduc

Teachers' interpersonal relationships and instructional expertise: How are they related?

R.M. van der Lans[a,b,*], J. Cremers[c,d], I. Klugkist[c], R. Zwart[e]

[a] Department of Teacher Education, Faculty of Social and Behavioral Sciences, University of Groningen, the Netherlands
[b] Curium-LUMC, Department of Child and Adolescent Psychiatry, Leiden University Medical Center, Leiden, the Netherlands
[c] Department of Methods and Statistics, Utrecht University, Utrecht, the Netherlands
[d] Biostatistics, Department of Public Health, University of Copenhagen, Denmark
[e] Department of Education, Utrecht University, Utrecht, the Netherlands

ARTICLE INFO

Keywords:
Teacher expertise
Interpersonal relationships
Student perceptions
Teacher-student relationships
Teaching quality

ABSTRACT

This study explores new directions to study and combine measurements of instructional expertise and teachers' interpersonal relationships with students. The sample comprises 34 in-service teachers. The My Teacher questionnaire (MTQ) was used to operationalize teachers' instructional expertise. The Questionnaire of Teacher Interaction (QTI) was used to describe teachers' interpersonal relationships with students. Hypotheses were tested using circular mixed-effects models. Results indicate that teachers' interpersonal relationships differ at successive levels of instructional expertise. Results further indicated that increases in instructional expertise are associated with a sharp decrease in the within-class variance in interpersonal relationships. Specifically, the higher teachers' instructional expertise, the more teachers' interpersonal relationships are described as "directing" and "helpful" by *all* students in the class.

1. Introduction

Expert teachers are characterized by expertise in explaining and helping students to understand core concepts, i.e. instructional expertise, and by expertise in maintaining productive interpersonal relationships with students, i.e. interpersonal expertise (Cohen & Goldhaber, 2016). Several studies address associations between these two characteristics of teaching quality (e.g., Downer, Stuhlman, Schweig, Martínez, & Ruzek, 2015; Ferguson & Danielson, 2014; Pianta & Hamre, 2009). In these studies, the concepts instructional and interpersonal expertise are part of the broader construct teacher expertise, which becomes evident in the quality of teaching. Although, studies generally distinguish between interpersonal and instructional expertise, studies routinely adminster questionnaires that mix items related to both concepts and typically apply the same psychometric procedures to these items. This study applies the Interpersonal Circle of the Teacher (IPC-T) to operationalize teachers' interpersonal expertise (Wubbels, Brekelmans, Brok, & Tartwijk, 2006) which only purpose is to describe teachers' interpersonal relationships with students. To operationalize teachers' instructional expertise, the study applies the international comparative analysis of learning and teaching (ICALT) model (Maulana, Helms-Lorenz, & Van de Grift, 2015; Van de Grift, Helms-Lorenz, & Maulana, 2014). Although the ICALT model is not uniquely focused on instructional expertise, it is dominated by items tapping teachers' instructional behavior. The ICALT model distinguishes six hierarchical levels of expertise of which only one level is related to interpersonal relationships.

Another reason to choose the ICALT model, is its connection with theory and models describing development of instructional expertise, including Antoniou and Kyriakides (2013); Berliner (2004); Ericsson (2006); Huberman (1993); Kagan (1992); Kyriakides, Creemers, and Antaniou (2009); Van de Grift et al. (2014) and van der Lans, van de Grift and van Veen (2015, 2017, 2019). The above studies are arguably different in various respects, but all understand the development of instructional expertise as a continuous progressive process in which beginning teachers start with learning basic teaching skills and in which expert teachers are characterized by mastering all teaching skills. Descriptions of this continuous progression show considerable overlap between studies (e.g., van der Lans et al., 2017). For example, most regard classroom management and structuring explanations as basic teaching skills that likely develop already in the beginning phases of teaching (e.g., Berliner, 2004; Kyriakides et al., 2009). However, results are inconsistent about the position of interpersonal relationships on this continuous progression. Results in Fuller (1969) and Van de Grift et al.

* Corresponding author at: Curium-LUMC, Leiden University, Endegeesterstraatweg 27, 2342 AK, Oegstgeest, the Netherlands.
E-mail address: r.m.van_der_lans@curium.nl (R.M. van der Lans).

https://doi.org/10.1016/j.stueduc.2020.100902
Received 8 June 2019; Received in revised form 20 June 2020; Accepted 22 June 2020
Available online 02 July 2020
0191-491X/

(2014) indicate that development of expertise in establishing interpersonal relationships concentrates in the beginning and novice phases of teaching. Results in Kyriakides et al. (2009), instead, indicate that the development of expertise in establishing interpersonal relationships extends beyond the phase of novice teaching. This difference might reflect the use of distinct and perhaps incomplete conceptualizations of interpersonal relationships and interpersonal expertise. By using the IPC-T, the study can add an unique and broad picture of teachers' interpersonal relationships. This can potentially add further understanding about the role of interpersonal relationships and interpersonal expertise in models describing the development of instructional expertise. Accordingly, the study's research question is: *How do teachers' relationships with students differ for teachers having distinct level of instructional expertise?*

2. Theoretical background

2.1. Instructional expertise and the ICALT model

In this study, teachers' instructional expertise is measured by observable behaviors, initiated by teachers, and of which empirical results indicate that they help to establish a learning environment that maximizes students' learning opportunities. This conceptualization overlaps with conceptualizations found in literature of teacher effectiveness (e.g., Hattie, 2009; Kyriakides, 2013; Muijs et al., 2014; Van de Grift et al., 2014). Studies on teacher effectiveness typically cluster the various observable teaching behaviors into fewer factors or domains. The current consensus suggests that the diversity in teachers' classroom practices is well summarized by five to seven clusters (e.g., Bell, Dobbelaer, Klette, & Visscher, 2019; Muijs et al., 2014). The ICALT model also summarizes teachers' classroom practices using six clusters, namely: safe and stimulating learning climate, efficient classroom management, clear and structured explanation, intensive and activating instruction, teaching learning strategies, and differentiation (Maulana et al., 2015; Van de Grift et al., 2014). In addition, the ICALT model orders these six clusters hierarchically on a single progressive continuum of instructional expertise (Van de Grift et al., 2014; van der Lans et al., 2015, 2019). This continuum starts with least the complex cluster of teaching behaviors, which are learned by novice teachers, and ends with the most complex cluster of teaching behaviors, which are only mastered by expert teachers. In the ICALT model, the term "complexity" is connected to an idea of conditionality. Hence, high complexity indicates that effective performance is conditional on various other skills, but not that the skill itself is difficult or complex to perform. We will describe each of the six ICALT levels briefly from least to most complex.

Level 1. *Establishing a safe learning climate*. A safe learning climate reflects an overall trust and respect among students and between students and teacher. Teachers operating at this level can still experience problems with establishing mutual respect and trust between teacher and students and among students. Students may regard the classroom as unsafe and refocus attention on strategies to guard themselves instead of listening and learning from the teacher. Studies show the importance of trust and respect for student learning (e.g., Pianta & Hamre, 2009; Wentzel, 2002).

Level 2. *Establishing an efficient classroom management*. Teachers' classroom management is qualified as efficient when the teacher succeeds to establish procedures, routines, and rules about where and how learning takes place (e.g, Korpershoek, Harms, de Boer, van Kuijk, & Doolaard, 2016; Muijs & Reynolds, 2003). Teachers operating at this level have established a basic level trust between teacher and student and among students. They experience problems with their classroom management. They lose time when shifting between tasks or assignments. Also, students do not always know what they should do or where to find required materials.

Level 3. *Establishing clear and structured explanations*. Clear and structured explanations prompt students' prior knowledge, emphasize critical knowledge, and check students' comprehension of content (e.g., Muijs & Reynolds, 2003; Rosenshine, 1995). Teachers operating at this level of expertise have established a level trust between teacher and student and among students and their efficiency in classroom management allows them to provide clear and structured front-class explanations. The teacher is acquiring skill in providing clear instructions. Typically, instructions are teacher-centered. In terms of Bloom's taxonomy of learning objectives, teachers teaching at this level of instructional expertise aid students to comprehend and understand (Krathwohl, 2002).

Level 4. *Establishing an intensive and activating instruction*. Intensive and activating instructions stimulate teacher–student and student–-student interaction by questioning, collaborative group work, having students explain topics to one another, or asking students to think aloud (Abrami et al., 2015; Muijs & Reynolds, 2003). Teachers operating at this level have acquired the basic skills to instruct students clearly. They are now are exploring different ways to organize student learning around specific assignments, questions, problems or tasks. Teacher-student interactions are longer and aim to stimulate students to apply the learned material to novel problems and/or to connect the learning material to other previously learned knowledge. Teaching behaviors related to intensive and activating instruction are a means to achieve the learning objectives of application and elaboration (Krathwohl, 2002).

Level 5. *Teaching students learning strategies*. Learning strategies enhance students' metacognitive skills and self-regulated learning (Abrami et al., 2015). Teachers operating at this level have acquired basic skills to instruct students clearly and have acquired different routines to organize student learning around assignments, questions, problems or tasks. These routines enable them to use a minimal amount of front-class instruction and more intensive interactions with and between students. Teachers are exploring ways to include process feedback within their interactions with students. Instructional behaviors related to teaching learning strategies are a means to achieve the learning objectives of synthesizing and evaluation (Krathwohl, 2002).

Level 6. *Learning how to differentiate instruction.* Differentiation means that teachers adjust their instructional practice to specific students' learning needs by, for example, allowing flexible time to complete assignments or providing additional explanation to small groups (e.g., Reis, McCoach, Little, Muller, & Kaniskan, 2011). Teachers operating at this level have acquired different routines to instruct and can organize student learning around specific tasks. The process feedback provides them with rich information about individual differences between students. Teachers are learning how to use this information to adjust assignments, questions, problems or tasks to meet the individual needs of students.

Several studies operationalizing the ICALT model report good fit of the hierarchical ordering in teaching behaviors with data gathered with student questionnaire and classroom observation instruments (Maulana et al., 2015; Van de Grift et al., 2014; van der Lans et al., 2015, 2018, 2019). These studies show that: 1) the above hierarchical ordering can be found in samples of teachers with different years of experience (i.e. samples of student teachers, beginning teachers and experienced teachers); 2) that the ordering is found with classroom observation and student questionnaire methods, and 3) that the ordering broadly parallels Fuller's (1969) and Berliner's (2004) descriptions of development in instructional expertise. Fig. 1 visualizes the hierarchical ordering of the six levels. Note that the ordering in levels is meant to assign meaning and facilitate interpretation. In reality, the boundaries between 'levels' are fluid meaning that, for example, items describing teaching behaviors related to the level 'efficient classroom management' mix with items related to 'safe learning climate' (at the lower side) and mix with items related to 'clear and structured explanations' (at the higher side) (e.g., see the Table 2 in van der Lans et al., 2019). Fig. 1 also visualizes connections with Berliner's (2004) and Fuller's

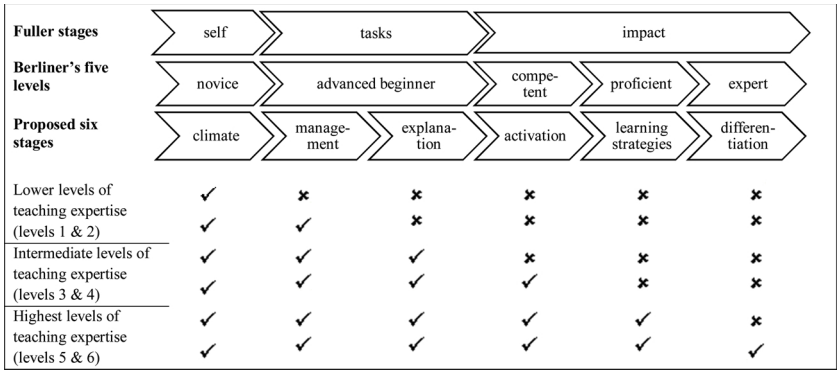

Fig. 1. The six levels of teachers' instructional expertise and their connection to Fuller's three stages and Berliner's five sequential phases of teaching expertise. For detailed descriptions of the overlap with Fuller's stages and Berliners levels of expertise it is referred to Van der Lans et al. (2017).
Notes: Check marks indicate that the teaching behaviors associated with this level likely are successfully performed; crosses indicate that the behaviors likely are not successfully performed.

(1969) theories of development in instructional expertise. These connections are not detailed here. Interested readers are referred to van der Lans et al. (2017).

The ICALT model views the six levels as hierarchical, meaning that it hypothesizes that expertise in the first level is required to obtain expertise in the next level. Based on this hypothesis, the ICALT model proposes to use standardized observation and questionnaire instruments to identify teachers' current level of expertise and subsequently support the (deliberate) practice and training of skills associated with the next higher level of expertise (Van de Grift et al., 2014; van der Lans et al., 2015). This interpretation is consistent with a stage-interpretation of the hierarchical ordering. The validity of this stage-interpretation has, however, not yet received thorough empirical attention. Other models on instructional expertise, most notably Berliner's (2004) model, have proposed a phases-interpretation. In the phases-interpretation, teachers can be grouped according to level of expertise at any moment in time, but their current level of expertise has no implications for what teachers best can learn next. The difference between the phases- or stages-interpretation has substantial consequences for how models of expertise are used by coaches and trainers to support teachers' professional development, but has limited consequences for the operationalization and/or description of the continuous progression in instructional expertise. Both the phases- and stages- interpretation suggest a continuous and additive relationship between the observed teaching behaviors. This study applies psychometric models that are consistent with this suggestion and, hence, results can apply to the phases- and stages interpretation.

2.2. Teachers' interpersonal relationships

In this study, teachers' interpersonal relationships is measured using the Interpersonal Circle for the Teacher (IPC-T) (previously known as the Model of Interpersonal Teacher Behavior (MITB)) (Wubbels et al., 2006). The IPC-T describes how students typify their relationship with their teachers using a circular scale. The circular scale is usually subdivided into eight octants which represent eight types of interpersonal relationships, namely: directing (locations 0° to 45°), helpful (locations 45° to 90°), understanding (locations 90° to 135°), compliant (locations 135° to 180°), uncertain (locations 180° to 225°), dissatisfied (locations 225° to 270°), confrontational (locations 270° to 315°), and imposing (locations 315° to 360°) (see Fig. 2). Recently, a ninth type, struggling, was added that describes teachers whose class average score cannot clearly be assigned to any of the above octants (Pennings & Hollenstein, 2020). The name "struggling" was assigned to this ninth type because teachers assigned to it mostly show low levels of instructional expertise as observed by low efficiency of classroom management procedures.

The IPC-T is linked to a particular branch of psychometric models called circumplex models (Browne, 1992). The circular scale reflects the associations among the eight different types of teacher-student relationships (Van Tartwijk, Mainhard, Brekelmans, den Brok, & Levy, 2014). Specifically, interpersonal relationships located close to one another, for example, the relationship types "uncertain" and "dissatisfied", are positively correlated and interpersonal relationships located opposite to one another, for example the relationship types "directing" and "uncertain", are negatively correlated. Interpersonal relationships located at 90 degrees, for example the relationship types "directing" and "understanding", share no correlation.

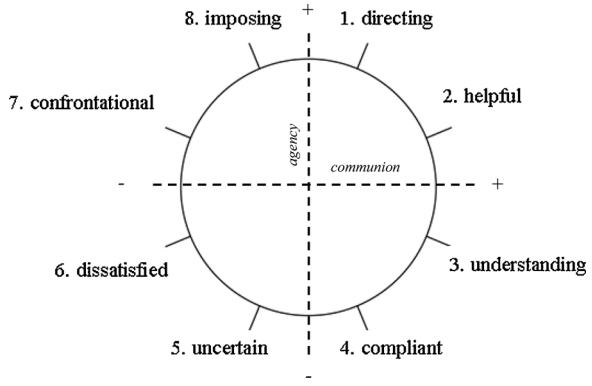

Fig. 2. IPC-T and the eight interpersonal relations. The two independent dimensions are added in dashed lines and the names applying to these dimensions are in italics.

As illustrated in Fig. 2, the circular scale lies in two-dimensional space. The two dimensions of this space are commonly referred as communion (or affiliation) and agency (or control) and are also visualized in Fig. 2. Communion reflects the overall warmth of teachers' communication towards the students and agency reflects the teachers' control over the students (Van Tartwijk et al., 2014; Wubbels et al., 2006). Previous studies frequently applied the sine and cosine functions to transform and locate the circular scores on these two linear dimensions (e.g., Mainhard, Brekelmans, den Brok, & Wubbels, 2011; Van Tartwijk et al., 2014; Wubbels et al., 2006). A description of this transformation is presented in the supplementary file (see page 6 "computation of the circular scores"). Because the sine and cosine are independent mathematical functions, the transformation defines communion and agency as two independent variables. This type of transformation is warranted by theory. Several psychological theories postulate that two independent dimensions suffice to describe interpersonal relationships (e.g., Leary, 1957; Pincus & Gurtman, 2006; Strack, 1996). Furthermore, the decomposition has proven practical utility and has successfully been applied in many prior studies both inside and outside educational sciences to communicate results and study properties of persons' interpersonal relationships (Pincus & Gurtman, 2006; Strack, 1996; Van Tartwijk et al., 2014; Wubbels et al., 2006). Nonetheless, a decomposition of the unidimensional circular scale into measurements on two independent dimensions also has drawbacks. Information about teachers' location on both dimensions is required to completely describe teachers' original circular score (Van Tartwijk et al., 2014). This is also visible in Fig. 2, where it can be

observed that teachers located within the octants "imposing" and "uncertain" have similar location on the dimension communion. Univariate analyses on the single dimension communion would ignore such differences. Multivariate analyses can be used to include the two dimensions simultaneously. However, given the absence of linear dependency, the standard multivariate analysis will not make any adjustments.

The present study applies a novel statistical approach, specifically it will apply circular mixed-effects models (Nuñez-Antonio & Gutiérrez-Peña, 2014). This novel approach is rooted in circular statistics and its key advantage is that it enables a direct analysis on the circular scores. Cremers, Mainhard and Klugkist (2019) give a detailed outline of this approach and the advantages that it has for analyzing data from the interpersonal circumplex. To communicate results, we will use the predicted teachers' location on the circumference, where, for example, teachers located between 0 to 45 degrees are reported to have a "directing" type of interpersonal relationship with their students. This novel approach aligns with recent studies that used the eight octants to distinguish between interpersonal relationship styles (e.g., Van der Want et al., 2015, 2018). However, it departs from other studies that also apply the term interpersonal relationship styles (e.g., Wubbels & Brekelmans, 2005; Wubbels et al., 2006). The latter mentioned studies use a slightly different taxonomy constructed based on the two dimensions communion and agency.

2.3. Prior findings related instructional expertise and interpersonal relationships

Prior findings indicate that few students describe their interpersonal relationship with the teacher as "confrontational" or "dissatisfied". To illustrate, Wubbels et al. (2006) mention that seven out of the 197 teachers were by the class average score typified as "confrontational" or "dissatisfied". Wubbels and Brekelmans (2005) report the number 68 out of 501 teachers (see their Table 3).[1] These results align with findings of the ICALT, which suggest that only few teachers fail to establish a basic level of trust and respect among students (Maulana et al., 2015; Van de Grift et al., 2014; van der Lans et al., 2015, 2018).

Wubbels et al. (2006) report that teachers' agency increases during the first seven years of teaching, but also that there are no substantial changes in teachers' communion. In terms of the relationship, results indicate that a larger portion of the beginning and inexperienced teachers are perceived by students as low in agency, i.e. student typify teachers as "uncertain", "compliant" or "understanding". Of the teachers with 6–10 years' experience are a larger portion is perceived by students as "directing" and "helpful". Near the end of the career, teachers are more often characterized by students as "imposing" (Wubbels et al., 2006). Cross-sectional studies on teachers' instructional expertise, as operationalized by the ICALT model, report an increase in expertise during the first 10 years of experience (Van de Grift, Van der Wal, & Torenbeek, 2011). Furthermore, instructional expertise slightly decreases near the end of the teaching career. Combined, these results give the impression that higher levels of instructional expertise are associated with a higher tendency of teachers to be described by the class (average) as: "directing", and "helpful".

Two prior studies examined the association between interpersonal relationships and instructional expertise directly, namely Brekelmans, Sleegers, and Fraser (2001) and den Brok (2001). Both these studies divided their samples into teachers who apply intensive and interactive instructions and those who do not apply interactive instructions and report differences in teachers' interpersonal relationships between these two groups. According to Brekelmans et al. (2001), teachers applying interactive instructions more frequently had "directing", "helpful", "understanding" types of relationships compared to teachers not applying interactive instructions. The effect size is large (eta^2 = .64), suggesting the association is worthy of further exploration. Contrary to the impression that stemmed from the cross-sectional studies, the studies by Brekelmans et al. (2001) and den Brok (2001) report that the use of interactive instructions also coincides with higher scores on the communion dimension. This indicates that increases in instructional expertise may not be limited to a shift on the agency dimension. Finally, like the present study, these two prior studies examined the association between perceived interpersonal relationships and teachers' instructional expertise. By using the ICALT hierarchy of six levels in instructional expertise, the present study can replicate, complement and further detail these prior findings. In specific, the broader dichotomy "no interactive instruction" (i.e. ICALT levels 1, 2, 3) and "interactive instruction" (i.e. ICALT levels 4, 5, 6) is in this study further subdivided into six separate levels of instructional expertise.

2.4. This study

Based on the above, the current study explores whether teachers' interpersonal relationships differ at successive levels of instructional expertise. The study examines evidence related to two hypotheses:

Hypothesis 1. Teachers' interpersonal relationships with students varies for the teachers having distinct levels of instructional expertise

Fig. 3 provides an example pattern that would support hypothesis 1. The Figure depicts the six hierarchical levels and for the benefit of visualization it clusters the lowest, intermediate and highest two levels. At the right side three hypothetical IPC-T circles are illustrated; one related to the lowest two levels of instructional expertise, one related to the intermediate two levels, and one related to the highest two levels. The three circles illustrate one possible direction of how the interpersonal relationships may differ at successive levels of instructional expertise.

In addition, with newly proposed circular mixed-effects model (Cremers et al., 2019) we obtain a means to explore whether the variance in the student perceived interpersonal relationships differs at successive levels of instructional expertise. An initial quick scan of the data showed lower variance for teachers at higher levels of instructional expertise. Therefore, we also decided to test whether variance in student perceived interpersonal relationships decreases with increasing levels of instructional expertise. Formally stated:

Hypothesis 2. Variance in teachers' interpersonal relationships with students decreases at successive levels of instructional expertise.

Fig. 4 provides an example pattern that would support hypothesis 2. The Figure again depicts the six hierarchical levels and at the right side three IPC-T circles; one related to the lowest two levels of instructional expertise, one related to the intermediate two levels, and one related to the highest two levels.

3. Method

This research was approved by the board of the Authors' University department as being in accord with the principles and ethics of human subject research. School and teacher participation in the project were voluntary, and participating schools received no funding.

3.1. Sample and data

The data were obtained as part of two projects: one school participated in a national induction project; the other participated in a smaller PhD research project. The total sample comprised 997 student questionnaires related to 39 teachers. Of these 997 questionnaires, the 816 were eligible for this study. The criteria for inclusion were: (1)

[1] Note that these studies applied a slightly deviating taxonomy. The mentioned numbers refer to the number of teachers classified as "drudging", "repressive" and "uncertain/aggressive".

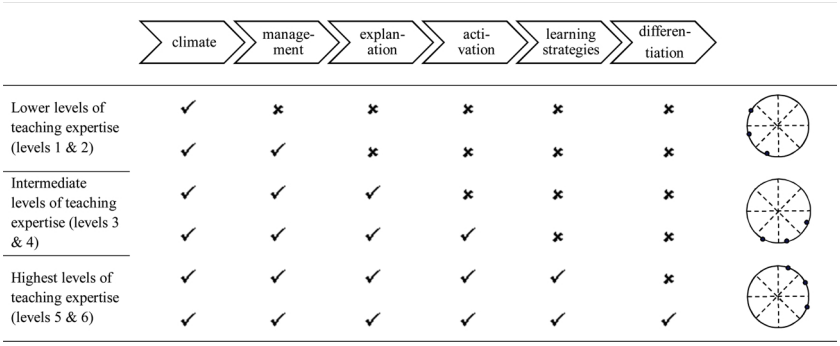

Fig. 3. Visualization of H1: Differences in teachers' type of interpersonal relationship when teachers' level of instructional expertise increases.

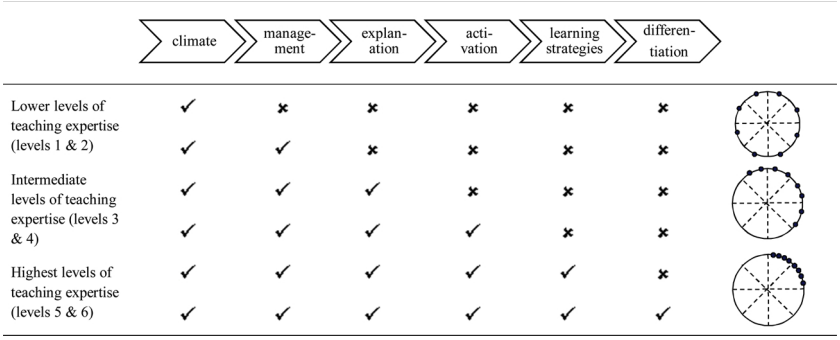

Fig. 4. Visualization of the decrease in variance of teachers' interpersonal relations for increasing levels of teachers' instructional expertise.

questionnaires should count no more than 4 missing values ($n_{omitted}$ = 99) and (2) questionnaires involve subject domains other than: mentoring, physical education and visual arts $n_{omitted}$ = 92. Note that the second criterion led to omitting all student data related to five teachers. Students attended Dutch secondary school (grades 1–4),[2] student age ranged between 12 and 17 years, and 50.6 % are girls. The eligible 816 questionnaires assess 34 in-service teachers, teaching various subjects including biology, geography, German, English, history, physics, Dutch, and math. The teachers' experience ranged between 4 and 31 years, with an average of 16 years.

3.2. Procedures

Two student questionnaires were administered, one measuring teachers' instructional expertise (the My Teacher questionnaire (MTQ)) and one measuring teachers' interpersonal relationships (the Questionnaire on Teacher Interaction (QTI)). Questionnaire administration was not counterbalanced. All students first responded to the questionnaire measuring instructional expertise and then responded to the questionnaire measuring interpersonal relationships. Questionnaires were administered by the teachers as part of the school's standard evaluation and feedback procedure. The time frame of administration spanned seven months, from December 2016 to June 2017. However, the timing of survey administration has shown to have little effect on the survey outcomes (e.g., Benton & Cashin, 2012; Mainhard et al., 2011).

3.3. Measures and operationalization

3.3.1. Instructional expertise

The My Teacher questionnaire (MTQ) is applied to operationalize teachers' instructional expertise. The MTQ is the student version of the ICALT observation instrument (Maulana & Helms-Lorenz, 2016; van der Lans et al., 2019). Several studies corroborate its reliability and validity (Maulana et al., 2015; van der Lans et al., 2015, 2019). Also, research reports that the ICALT and MTQ order the six levels of instructional expertise similarly except for the final two levels (van der Lans et al., 2019). Within the MTQ the order of the fifth (learning strategies) and sixth (differentiation) levels tend to be blurred or reversed compared to the order found with the ICALT observation instrument. Note that the order described in the background is based on the observation instrument.

The administered version of the MTQ consists of 40 items describing teaching behaviors related to the six domains, such as "My teacher makes sure that I use my time effectively" (efficient classroom management) or "My teacher asks me how I am going to learn the content of the lesson" (teaching learning strategies). This version of the MTQ reflects the findings of two previous validation studies (Maulana et al., 2015; van der Lans et al., 2015). However, not all items included have been validated to fit the expected item response pattern. Therefore, the validity of these items was assessed using Rasch (1960) and Mokken (1971) model assumption tests, as follows:

1 *Differential item functioning.* The two schools participated in two projects that used two slightly different versions of the MTQ. One school participating in the PhD projects had students rate 40-items on two categories, whereas the other school used the 41-item version with four categories (which was then dichotomized). Given these small differences, we checked for differential item functioning (DIF) between the two types of questionnaires.
2 *Local independence.* The assumption of the local independence was assessed using Ponocny (2001) T_1 and T_{1m}.
3 *Non-intersection of item response functions.* We evaluated whether item response functions intersected or not (Van der Ark, 2007).

With these steps, we identified 29 items that fit the assumptions. Internal consistency of the selected items was good: Cronbach's alpha = 0.89, the minimum split-half reliability = 0.82. The complete results of the validation study are available in a supplementary file

[2] In the United States, these grades are equivalent to grades 7–10.

which also provides information about the final ordering of the levels in instructional expertise.

3.3.2. Interpersonal relationships

Teachers' interpersonal relationships with students was operationalized according to the Questionnaire on Teacher Interaction (QTI; Wubbels et al., 2006). There are several versions of the QTI. This version consisted of 24 items, scored on a 5-point Likert scale. The Likert scale data was calibrated using a circumplex model (Browne, 1992; Grassi, Luccio, & Di Blas, 2010). The fit of the circumplex model to the data was adequate (χ^2(df = 226, n = 816) = 820.36, p = 0.00; CFI = 0.94, TLI = 0.92; root mean error of approximation (RMSEA) = 0.057 [90 % confidence interval = 0.053, 0.061]). The supplementary file contains information about the empirical octant positions on the circumference.

The circumplex model output provides a circular ordering of items in angular positions (in degrees), but currently it does not provide empirical estimates of person locations on the circumference. A transformation had to be applied to estimate student locations on the circumference. This transformation requires two steps. First, the scores are decomposed into measurement on the two linear dimensions: communion and agency. This step applies the cosine and sine transformation that was already discussed in the background section. Second the inverse of the tangent was used to transform students' communion and agency scores back onto the circumference (the supplementary file provides exact details).

3.3.3. Descriptive statistics

Table 1 contains the average QTI scores and variance. The variance of a circular variable such as in this study the QTI runs from 0.00 to 1.00, where 0.00 means that all teachers' scores are located at the same point on the circumference and 1.00 means that teachers' scores are uniformly distributed around the circle (Fisher, 1995). The Table reports averages and variance on the QTI for teachers grouped according to their class mean level of instructional expertise (aggregated teacher level) and for students grouped by their unique experience of teachers' level of instructional expertise (disaggregated student level). Teachers' class average level of instructional expertise varies less compared to students' unique experiences of instructional expertise. In this specific sample, no teachers had class average MTQ scores indicative of level 1 ("safe learning climate") or level 6 (which in the MTQ is "teaching students learning strategies"). Yet, as Table 1 shows, individual students may experience their teacher's instructional expertise as being of level 6 (or level 1). These students tend to experience their teacher's instructional expertise better (or worse) than most peer-classmates.

Table 1
Descriptive statistics for the Questionnaire of Teacher Interaction (QTI), grouped by the six levels of expertise. The $n_{(teachers)}$ column indicates the number of teachers whose class average score falls within that level of expertise. The $n_{(students)}$ column indicates the number of students that experience their teachers' teaching within that level of instructional expertise.

$n_{(teachers)}$	$n_{(students)}$	"My Teacher" Stage	Teachers' Average (Variance)	Students' Average (Variance)
–	143	Safe learning climate	–	354.07° (0.62)
3	41	Efficient classroom management	353.12° (0.45)	14.56° (0.25)
9	148	Clarity of instruction	21.94° (0.28)	21.47° (0.16)
16	193	Activating teaching methods	25.95° (0.08)	25.24° (0.06)
6	223	Differentiation	31.72° (0.01)	29.56° (0.01)
–	68	Learning strategies	–	32.34° (0.04)

Note. An "- "indicates that no teacher's is scored within that level of expertise. Thus, all teachers were by the class average student experienced as able to establish a basic level of respect in the classroom. Also, no teachers were by the class-average student experienced as skilled in all aspects of instruction including the most complex aspect: teaching learning strategies.

3.4. Model and analysis plan

The analysis strategy intended to quantify the evidence in favor of the hypotheses at the student- and teacher-level. Hypotheses are examined using the parameters estimated within a set of Bayesian projected normal circular mixed-effects models (Nuñez-Antonio & Gutiérrez-Peña, 2014). In these models, the circular variable teachers' interpersonal relationship is the dependent variable. Four competing models were compared using various fit indices to verify which model gave the most accurate parameter estimates. Results presented are related to the best fitting model. Details about the model comparison are found in the supplementary file.

3.4.1. Analysis of hypothesis 1

To examine whether: *Teachers' interpersonal relationships with students varies for the teachers having distinct levels of instructional expertise*, the ICALT scores were transformed into an ordinal score identifying teachers' level of expertise. The IPC-T circular scores were included as the dependent variable and the SAM, a circular regression coefficient (Cremers, Mulder, & Klugkist, 2018; Cremers et al., 2019), was used to examine mean differences in location on the circumference between the successive levels of instructional expertise. The hypothesis is examined at two levels simultaneously: one time using the class aggregate ratings of instructional expertise ($H1_{teacher}$) and one time using the student unique ratings of instructional expertise ($H1_{student}$). The formal hypotheses are:

$$\mathbf{H1_{teacher}}: \beta_2 MTQ\ Teacher_i \neq 0$$

$$\mathbf{H1_{student}}: \beta_1 MTQ\ Student_{ij} \neq 0$$

At the aggregated level, the analysis examines whether teachers' class average interpersonal relationship differs between teachers located at different levels on the progressive continuum of instructional expertise (n = 34 teachers). At the disaggregated level, the analysis examines whether students' unique experience of their interpersonal relationship with the teacher differs between students that experience the level of their teachers' instructional expertise differently (n = 816 students). Note that at the disaggregated level ($H1_{student}$) students from different classes are grouped together based on their similar experienced level of instructional expertise.

3.4.2. Analysis of hypothesis 2

The hypothesis: *variance in teachers' interpersonal relationships with students decreases at successive levels of instructional expertise*, was also examined at two levels: the class aggregate instructional expertise score ($H2_{teacher}$) and student unique instructional expertise score ($H2_{student}$), or formally:

$$\mathbf{H2_{teacher}}: (Var_{ij} \mid \text{expertise level } 1_{class\ mean}) > (Var_{ij} \mid \text{expertise level } 2_{class\ mean}) > (Var_{ij} \mid \text{expertise level n}\ldots_{\ class\ mean})$$

$$\mathbf{H2_{student}}: (Var_{ij} \mid \text{expertise level } 1_{student}) > (Var_{ij} \mid \text{expertise level } 2_{student}) > (Var_{ij} \mid \text{expertise level n}\ldots_{\ student})$$

Var_{ij} represents the posterior mode of the variance of the QTI ratings. For example, in $H2_{teacher}$ the parameter "expertise level $2_{class\ mean}$", examines the variance in student experienced interpersonal relationships between students within the same class and that are taught by a teacher whose class mean instructional expertise is of level 2. Furthermore, the parameter "expertise level $2_{student}$" stated in hypothesis $H2_{student}$, examines the variance in the students' experienced interpersonal relationships within the group of students experiencing their teachers' instructional expertise of level 2. Note again that in $H2_{student}$ groups students from different classes together based on their

Table 2
Posterior estimates for the regression parameters.

Variable	Posterior Mode	Posterior SD	Lower Bound 95 % HPD	Upper Bound 95 % HPD
Type of questionnaire (Type)	−0.05	0.07	−0.16	0.10
Instructional expertise ($MTQ_{Teacher\ level}$)	0.27	0.22	0.19	0.36
Instructional expertise ($MTQ_{Student\ level}$)	0.14	0.07	0.12	0.16

Notes: Highest Posterior Density (HPD) is a Bayesian statistic comparable to the regular 95 % confidence interval.
Notes: information about the independent variable "Type of questionnaire" can be found in the supplementary file.

similar experienced level of instructional expertise.

Support for the hypotheses is quantified using Bayes factors. A Bayes factor is a ratio of the fit and complexity of a hypothesis. The fit is computed as the proportion of iterations of the Multiple Chain Monte Carlo (MCMC) sampler in which the hypothesis holds (see Klugkist, Laudy, & Hoijtink, 2005). The complexity suggests the number of alternatives that could falsify the hypothesis. The complexity of $H2_{student}$ equals 1/6! (i.e., the predicted hierarchical ordering of the six levels divided by all possible alternative hierarchical orderings). For $H2_{teacher}$, complexity equals 1/4! The latter is smaller, because for the aggregate class means there are no teachers at levels 1 and 6 (see Table 1). For both hypotheses we compute two Bayes factors, one quantifying evidence that teachers' interpersonal relationships differ for the empirically found successive six levels of instructional expertise: i.e. $BF_{hypothesis}$, and one quantifying evidence that teachers' interpersonal relationships differ for any other possible ordering in levels, i.e., $BF_{complement}$. The Bayes factor for the complement can in our case be computed as 1 – $BF_{hypothesis}$. The evidence in favor of a hypothesis is computed as the ratio of the two Bayes factors, $BF_{hypothesis}/BF_{complement}$. If this ratio falls between 0 and 1, alternative orderings are more likely than the one specified by the hypotheses. If it is larger than 1, the alternative orderings are less likely.

4. Results

The results for hypotheses $H1_{teacher}$ and $H1_{student}$ indicate that interpersonal relationships differ for distinct levels of instructional expertise (Table 2). The direction is counter-clockwise, indicating that a teacher scoring in the lowest level of instructional expertise is positioned at 231.39 degrees in the model, suggesting an "uncertain" type of relationship. When the level of instructional expertise increases, the class average teacher-student relationship shifts following the sequence "compliant", "understanding", and "helpful"; those who score at the highest levels of instructional expertise are predicted to have "directing" types of interpersonal relationships.

The results for hypotheses $H2_{teacher}$ and $H2_{student}$ indicated that the variance decreases at successive levels of instructional expertise. Fig. 5 illustrates this decrease in the variation in teachers' interpersonal relationships between four classes taught by four different teachers whose class-average scores indicate different levels of instructional expertise. Top left presents a typical plot for a teacher with a class average instructional expertise indicative of level 2, top right a typical plot for a teacher at level 3, bottom left a typical plot for a teacher at level 4, and bottom-right a typical plot for a teacher at level 5. As the plots show, the higher teachers' instructional expertise is, the greater number of students in the class experience their relationships with this teacher as "helpful" or "directing". Plots of all 34 teacher are included in the appendix 2.

This decrease in variance is also found when students of different teachers are grouped together based on their perceived level of instructional expertise. The group of students perceiving their teachers' instructional expertise as of the lowest level, i.e. "safe learning climate" had largest variance, more specifically the results suggest that this level of instructional expertise can coincide with any type of interpersonal relationship. The variance in interpersonal relationships decreased successively in every group of students that experience higher levels of instructional expertise. At the highest level, there were no meaningful differences in how students described their relationship with the teacher, i.e. they all characterized their relationship with the teacher as "directing" and "helpful". Again, the plots at the student level are included in the appendix 2. Within this sample of 34 teachers and 816 students, we found not a single deviation from this decrease in variance in all 5000 iterations (MCMC samples) of the model.[3]

5. Discussion

This study explores the association between teachers' interpersonal relationships with students and their level of instructional expertise. The results signal that the acquisition of high levels of instructional expertise coincides with "directing" and "helpful" types of interpersonal relationships. Furthermore, higher levels of instructional expertise were only observed when *all* students in the class experience the teacher as "directing" and "helpful". The strength of this pattern was exceptional; we found no exception on it at the class level (n = 34) and no exception at the student level (n = 816). The results indicate that an instructional expertise of level 1 and 2 can coincide with any type of student perceived interpersonal relationship. An instructional expertise of levels 3 and 4 coincides with an absence of students experiencing their interpersonal relationship with the teacher as "confrontational" and "dissatisfied" and (at level 4) "uncertain". An instructional expertise of levels 5 and 6 coincides with the complete classroom of students experiencing their interpersonal relationship with the teachers as "directing" and "helpful". This finding provides novel insights about the association between instructional expertise and interpersonal relationships.

5.1. Relationship with prior studies

Results replicate prior evidence that the application of activating instructional methods coincides with an increase in communion in student perceived interpersonal relationships and with "directing", "helpful", "understanding" types of relationships (Brekelmans et al., 2001; den Brok, 2001). Furthermore, results also corroborate the impression obtained from longitudinal and cross-sectional studies describing development in interpersonal relationships (Wubbels et al., 2006) and instructional expertise (Van de Grift et al., 2011), that increases in instructional expertise coincide with an increase in the agency dimension. But the methods applied in this study can further detail how class-average perceived interpersonal relationships change at successive levels of instructional expertise. The study operationalized six hierarchical levels of instructional expertise, namely safe learning climate (level 1), efficient classroom management (level 2), clear and structured explanations (level 3), interactive and activating instructions (level 4), teaching learning strategies (level 5) and differentiation (level

[3] There were no deviations meaning that $BF_{hypothesis} = 1$ and the $BF_{complement} = 0$. Technically, we cannot divide by zero. Therefore, we replaced $BF_{complement}$ by an infinitely small number. As a result, the estimated evidence in favor of the hypotheses becomes infinitely large.

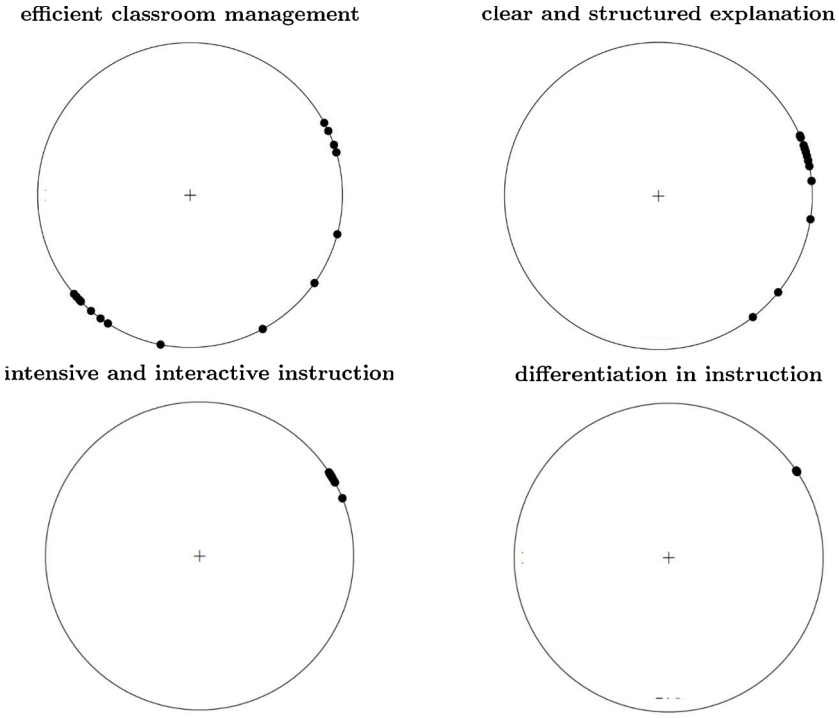

Fig. 5. Graphical illustration of the decrease in the within class-variance in student perceived interpersonal relationships at successively higher levels of instructional expertise. Every plot presents data of one 'typical' teacher. Note that plots for all 34 teachers are presented in the appendix 2 of the supplementary file.

6). At the lowest level of instructional expertise students likely describe their class-average interpersonal relationship as "uncertain". At subsequent levels of instructional expertise student descriptions are predicted to move over the circumference following the sequence "compliant", "understanding", "helpful", and finally "directing".

5.2. Implications for theory and practice

5.2.1. Insights and potential directions for research on instructional expertise

The ICALT model proposed by Van de Grift et al. (2011) deliberately applies a narrow conceptualization of interpersonal relationships. The logic behind this narrow conceptualization is that teachers should secure a basic level of trust and respect among students. This basic level of trust is sufficient to acquire instructional expertise and, thus, interpersonal expertise beyond this basic level is not strictly required. We note that the current results cannot completely falsify this logic. That is, the logic only claims that development of instructional expertise does not benefit from further acquisition in interpersonal expertise, which claim does not exclude the possibility that teachers keep investing in the acquisition of interpersonal expertise (as our results suggest). Hence, possibly the results only reflect that teachers' belief that their interpersonal relationships with students is important and, therefore, keep investing in improving them. Yet, we belief that the results of this study challenge the logic that only a basic level of trust is required. That is, if further development in expertise with establishing interpersonal relationships is not strictly required, then why have we not found any teacher with highest instructional expertise (i.e. level 5) in combination with types of interpersonal relationships other than "directing" and "helpful"? The question whether interpersonal expertise beyond some basic level is required to achieve instructional expertise is a pressing issue. Currently some teacher educators and schools are known to use theories like the here studied ICALT model of instructional expertise to allocate professional development resources. If so, they spend their resources likely on training and coaching of instructional expertise. However, based on the results of this study, it is possible that a subgroup of teachers may need investments in coaching and training on interpersonal expertise and, thus, *not* instructional expertise. Future research is advised to implement small scale experimental studies to examine the differential effectiveness of training in interpersonal relationships versus instructional expertise. Such studies also could further explore how to use the combination of interpersonal relationship and instructional expertise measures when training and coaching teachers.

The findings related to the second hypothesis were exceptionally strong. We speculate that this result possibly reflects another hierarchy possibly reflecting the development of expertise in interpersonal relationships which hierarchy unfolds parallel to the hierarchical levels of instructional expertise. The speculation suggests that the acquisition of successively higher levels of instructional expertise coincides with a successive decrease in the variation in student perceived interpersonal relationships. Again, large scale and longitudinal studies are required to adequately test this speculation.

5.2.2. Insights and potential future directions related to interpersonal relationships and the IPC-T

The exploration strengthens and adds to the interpretation of the ninth profile "struggling" reported by Pennings and Hollenstein (2020). Pennings and Hollenstein suggest that struggling teachers are characterized by (extremely) low levels of instructional expertise and disorganized classroom management, which in the present study are teachers having an ICALT instructional expertise of level of 1or 2. Our results suggest that classes of students taught by "struggling" teachers show high variability in student perceived interpersonal relationships. Due to this variability, struggling teachers may not have a consistent class average position on the circumference. Thus, although a sample estimated class-average position of "struggling" teachers can be estimated on the circumference, this estimated location is surrounded by a high level of uncertainty. When using circular scores, "struggling" teachers are identifiable by the large variance in student perceived interpersonal relationships. Summarized, the results corroborate the observation by Pennings and Hollenstein (2020) that a subgroup of

struggling teachers with low levels teaching proficiency has no clear class average position on the IPC-T.

This study applied circular-linear regression models to analyze the IPC-T model, which permit a direct analysis on the circular scores. We belief that this direct method has benefits over alternative indirect methods that decompose the circular scale into two measurements on linear dimensions or into eight measurements representing the octants. One demonstrated benefit are the visualizations. Visualization of circular outcomes has a direct interpretation to the theoretical IPC-T model. Furthermore, as discussed in the theoretical background, the circular mixed-effects models further help to decrease information loss. Nonetheless, the results also show that conclusions based on output from circular mixed-effects models closely mirror conclusions of previous studies that did not apply circular mixed-effects models (Brekelmans et al., 2001; den Brok, 2001; Wubbels et al., 2006). In sum, compared to the more traditional methods, circular-linear regression models further minimize data loss but also increase the required level of statistical expertise.

5.3. Limitations

We explored the association between teachers' level of instructional expertise and teachers' interpersonal relationships with student data. The study has several limitations, which can inform and improve future research on the topic. First, the study proposed to distinguish between the concepts: instructional expertise and interpersonal relationships. The exact operationalizations studied, however, still show minor overlap in conceptualization (i.e. the safe learning climate level). Another limitation pertains to the claims behind the applied model of instructional expertise. This model makes various claims about development in expertise. It is stressed that the study only examines associations between levels of instructional expertise and interpersonal relationships and that the study did not apply a cross-sectional or longitudinal design to support claims about development. Another third limitation pertains to the lack of counterbalanced procedure for the questionnaire administration. The students' perceptions of their relationship might have been shaped by their responses to questionnaire measuring instructional quality. Future research is advised to use counterbalanced administration procedures. Finally, the results rely on a select and small sample of 34 teachers in two schools. Partly, because of this limited sample size, the test of $H2_{teacher}$ was restricted to for four of the six levels. Replications are required to strengthen the results.

5.4. A final note

This study proposed that research on teaching can benefit from a strict separation between the concepts: (1) instructional expertise and (2) teachers' interpersonal expertise. Many of the contemporary instruments already mix items tapping interpersonal relationships and instructional expertise (Cohen & Goldhaber, 2016). The theories behind these instruments often detail differences between the two concepts, but apply the same psychometric method to items related to both concepts. For example, Downer et al. (2015) reports about a new student questionnaire composed of items tapping interpersonal relationships and instructional expertise and examines them jointly in the same factor analysis model. Almost never do studies seem to question whether a uniform psychometric approach fits best to all the underlying concepts. This study introduces two theoretically well-established lines of research, one clearly more associated with the concept instructional expertise, the other clearly more associated with interpersonal relationships. It applied distinct psychometric scaling procedures to these two concepts and innovative circular-linear mixed effects models to study associations between them. Although the pursued methods are highly innovative, we hope that this study may prove an inspiration to those seeking an alternative method to distinguish and study associations between these two key concepts of teacher expertise.

Author notes

Irene Klugkist and Jolien Cremers acknowledge financial support from a Vidi grant awarded to I. Klugkist from the Dutch Organization for Scientific Research (NWO 452-12-010).

Rikkert van der Lans acknowledges financial support from the Dutch ministry of Education (recipient Dr. Michelle Helms-Lorenz), grant number: OND/ODB-2013/45916 U).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.stueduc.2020.100902.

References

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research, 85*(2), 275–314. https://doi.org/10.3102/0034654314551063.

Antoniou, P., & Kyriakides, L. (2013). A dynamic integrated approach to teacher professional development: Impact and sustainability of the effect of improving teacher behavior and student outcomes. *Teaching and Teacher Education, 29*, 1–12. https://doi.org/10.1016/j.tate.2012.08.001.

Bell, C. A., Dobbelaer, M. J., Klette, K., & Visscher, A. (2019). Qualities of classroom observation systems. *School Effectiveness and School Improvement, 30*(1), 3–29.

Benton, S. L., & Cashin, W. E. (2012). *Student ratings of teaching: A summary of the research and literature. (IDEA Paper No. 50)*. Retrieved March 3, 2015, from http://www.ntid.rit.edu/sites/default/files/academic_affairs/Sumry%20of%20Res%20%2350%20Benton%202012.pdf.

Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society, 24*(3), 200–212.

Brekelmans, M., Sleegers, P., & Fraser, B. J. (2001). Teaching for active learning. In P. R. J. Simons, J. L. van der Linden, & T. Duffy (Eds.). *New learning* (pp. 227–242). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Browne, M. W. (1992). Circumplex models for correlation matrices. *Psychometrika, 57*(4), 469–497.

Cohen, J., & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observations. *Educational Researcher, 45*(6), 378–387.

Cremers, J., Mulder, K. T., & Klugkist, I. (2018). Circular interpretation of regression coefficients. *British Journal of Mathematical and Statistical Psychology, 71*(1), 75–95.

Cremers, J., Mainhard, T., & Klugkist, I. (2019). Assessing a Bayesian embedding approach to circular regression models. *Methodology*.

den Brok, P. (2001). *Teaching and student outcomes (Unpublished dissertation)*. Utrecht, The Netherlands: WCC.

Downer, J. T., Stuhlman, M., Schweig, J., Martínez, J. F., & Ruzek, E. (2015). Measuring effective teacher-student interactions from a student perspective: A multi-level analysis. *The Journal of Early Adolescence, 35*(5-6), 722–758.

Ericsson, K. A. (2006). *The influence of experience and deliberate practice on the development of superior expert performance. The Cambridge handbook of expertise and expert performance*, *38*, 685–705.

Ferguson, R. F., & Danielson, C. (2014). How framework for teaching and tripod 7Cs evidence distinguish key components of effective teaching. In T. J. Kane, & K. A. Kerr (Eds.). *Designing teacher evaluation systems* (pp. 98–143). San Francisco, CA: Wiley & Sons, Inc.

Fisher, N. I. (1995). *Statistical analysis of circular data.* Cambridge: Cambridge University Press.

Fuller, F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal, 6*, 207–226.

Grassi, M., Luccio, R., & Di Blas, L. (2010). CircE: An R implementation of Browne's circular stochastic process model. *Behavior Research Methods, 42*, 55–73. https://doi.org/10.3758/BRM.42.1.55.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* Abingdon, Oxon, UK: Routledge.

Huberman, M. (1993). *The lives of teachers.* New York, NY: Teachers College Press.

Kagan, D. M. (1992). Professional growth among pre-service and beginning teachers. *Review of Educational Research, 62*, 129–169.

Klugkist, I., Laudy, O., & Hoijtink, H. (2005). Inequality constrained analysis of variance: A bayesian approach. *Psychological Methods, 10*(4), 477–493.

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research, 86*(3), 643–680.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice, 41*(4), 212–218.

Kyriakides, L. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education, 36*, 143–152.

Kyriakides, L., Creemers, B. P. M., & Antaniou, P. (2009). Teacher behavior and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and Teacher Education, 25*, 12–23.

Leary, T. (1957). *An interpersonal diagnosis of personality*. New York: The Ronald Press.
Mainhard, M. T., Brekelmans, M., den Brok, P., & Wubbels, T. (2011). The development of the classroom social climate during the first months of the school year. *Contemporary Educational Psychology, 36*(3), 190–200.
Maulana, R., & Helms-Lorenz, M. (2016). Observations and student perceptions of the quality of preservice teachers' teaching behaviour: Construct representation and predictive quality. *Learning Environments Research, 19*(3), 335–357.
Maulana, R., Helms-Lorenz, M., & Van de Grift, W. (2015). Development and evaluation of a questionnaire measuring pre-service teachers' teaching behaviour: A Rasch modelling approach. *School Effectiveness and School Improvement, 26*(2), 169–194.
Mokken, R. J. (1971). *A theory and procedure of scale analysis: With applications in political research, Vol. 1*. The Hague, NL: Walter de Gruyter.
Muijs, D., & Reynolds, D. (2003). Student background and teacher effects on achievement and attainment in mathematics: A longitudinal study. *Educational Research and Evaluation, 9*(3), 289–314. https://doi.org/10.1076/edre.9.3.289.15571.
Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art – Teacher effectiveness and professional learning. *School Effectiveness and School Improvement, 25*(2), 231–256. https://doi.org/10.1080/09243453.2014.885451.
Nuñez-Antonio, G., & Gutiérrez-Peña, E. (2014). A Bayesian model for longitudinal circular data based on the projected normal distribution. *Computational Statistics & Data Analysis, 71*, 506–519. https://doi.org/10.1016/j.csda.2012.07.025.
Pennings, H. J., & Hollenstein, T. (2020). Teacher-student interactions and teacher interpersonal styles: A state space grid analysis. *The Journal of Experimental Education, 88*(3), 382–406.
Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*(2), 109–119.
Pincus, A. L., & Gurtman, M. B. (2006). interpersonal theory and the interpersonal circumplex: Evolving perspectives on normal and abnormal personality. In S. Strack (Ed.). *Differentiating normal and abnormal personality* (pp. 83–111). New York, NY, US: Springer Publishing Co.
Ponocny, I. (2001). Non-parametric goodness-of-fit tests for the Rasch model. *Psychometrika, 66*, 437–460.
Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Copenhagen: Nielsen & Lydiche.
Reis, S. M., McCoach, D. B., Little, C. A., Muller, L. M., & Kaniskan, R. B. (2011). The effects of differentiated instruction and enrichment pedagogy on reading achievement in five elementary schools. *American Educational Research Journal, 48*(2), 462–501.
Rosenshine, B. (1995). Advances in research on instruction. *The Journal of Educational Research, 88*(5), 262–268.
Strack, S. (1996). Introduction to the special series-interpersonal theory and the interpersonal circumplex: Timothy Leary's legacy. *Journal of Personality Assessment, 66*(2), 212–216.
van der Lans, R. M., van de Grift, W. J., & van Veen, K. (2015). Developing a teacher evaluation instrument to provide formative feedback using student ratings of teaching acts. *Educational Measurement: Issues and Practice, 34*(3), 18–27.
van der Lans, R. M., van de Grift, W. J., & van Veen, K. (2017). Individual differences in teacher development: an exploration of the applicability of a stage model to assess individual teachers. *Learning and Individual Differences, 58*, 46–55.
van der Lans, R. M., van de Grift, W. J., & Van Veen, K. (2018). Developing an instrument for teacher feedback: Using the Rasch model to explore teachers' development of effective teaching strategies and behaviors. *The Journal of Experimental Education, 86*(2), 247–264.
van der Lans, R. M., van de Grift, W. J., & van Veen, K. (2019). Same, similar, or something completely different? Calibrating student surveys and classroom observations of teaching quality onto a common metric. *Educational Measurement: Issues and Practice, 38*(3), 55–64.
Van de Grift, W. J. C. M., 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. https://doi.org/10.1016/j.stueduc.2014.09.003.
Van de Grift, W. J. C. M., Van der Wal, M., & Torenbeek, M. (2011). Ontwikkeling in de pedagogische didactische vaardigheid van leraren in het basisonderwijs. [Primary teachers' development of pedagogical didactical skill]. *Pedagogische Studiën, 88*, 416–432.
Van der Ark, A. L. (2007). Mokken scale analysis in R. *Journal of Statistical Software, 20*(11), 1–19. https://doi.org/10.18637/jss.v020.i11.
Van der Want, A. C., den Brok, P., Beijaard, D., Brekelmans, M., Claessens, L. C., & Pennings, H. J. (2015). Teachers' interpersonal role identity. *Scandinavian Journal of Educational Research, 59*(4), 424–442.
Van der Want, A. C., Den Brok, P., Beijaard, D., Brekelmans, M., Claessens, L. C., & Pennings, H. J. (2018). Changes over time in teachers' interpersonal role identity. *Research Papers in Education, 33*(3), 354–374.
Van Tartwijk, J. W. F., Mainhard, M. T., Brekelmans, M., den Brok, P. J., & Levy, J. (2014). Docent-leerling interacties en het sociaal klimaat in de klas [teacher-student interactions and the classroom climate]. In J. Van Tartwijk, M. Brekelmans, P. Den Brok, & T. Mainhard (Eds.). *Theorie en praktijk van leren en de leraar: Liber Amicorum Theo Wubbels* (pp. 25–43). Amsterdam: Uitgeverij SWP.
Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development, 73*(1), 287–301.
Wubbels, T., & Brekelmans, M. (2005). Two decades of research on teacher–student relationships in class. *International Journal of Educational Research, 43*(1–2), 6–24.
Wubbels, T., Brekelmans, M., Brok, P., & Tartwijk, J. (2006). An interpersonal perspective on classroom management in secondary classrooms in the Netherlands. In C. Evertson, & C. Weinstein (Eds.). *2006 Handbook of classroom management: Research, practice, and contemporary issues* (pp. 1161–1191). Mahwah, NJ: Lawrence Erlbaum Associates.