



OPEN ACCESS

EDITED BY

Aldo Bazán-Ramírez,
Universidad Nacional José María Arguedas,
Peru

REVIEWED BY

Walter Capa,
National University Federico Villareal, Peru
África Borges,
University of La Laguna, Spain

*CORRESPONDENCE

Carina Wikman
✉ carina.wikman@specped.su.se

RECEIVED 23 August 2024

ACCEPTED 06 November 2024

PUBLISHED 11 December 2024

CITATION

Wikman C, Allodi MW and Ringer N (2024)
Classroom observations in elementary
schools with Educational Climate Assessment
Tool (EduCAT): the concurrent validity with
Classroom Assessment Scoring System.
Front. Educ. 9:1485260.
doi: 10.3389/educ.2024.1485260

COPYRIGHT

© 2024 Wikman, Allodi and Ringer. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Classroom observations in elementary schools with Educational Climate Assessment Tool (EduCAT): the concurrent validity with Classroom Assessment Scoring System

Carina Wikman*, Mara Westling Allodi and Noam Ringer

Department of Special Education, Stockholm University, Stockholm, Sweden

The classroom social–emotional climate is critical to students' wellbeing and learning ability. In practice and research, an assessment instrument with strong validity is invaluable, providing crucial information about the classroom's social–emotional and learning climate. This allows for a more accurate appreciation of the quality of the learning environment. A valid instrument is helpful and essential in improving the classroom climate in practice and demonstrating intervention outcomes in research. The primary aim was to evaluate the concurrent validity of the Educational Climate Assessment Tool (EduCAT)—a newly developed classroom observation instrument based on a solid theoretical framework and previous empirical research. The concurrent validity testing involved examining pairwise and multivariate canonical correlations between scores from the EduCAT scales (stimulation, improvement, structure, relationship, and influence) and scores from the Classroom Assessment Scoring System (CLASS®) administered concurrently in 30 kindergarten to third grade (K-3) classrooms in four schools in Sweden. The pairwise correlation results show that the scores of the two observation instruments have strong and moderate positive correlations, as hypothesized. The canonical correlation identifies which scales in EduCAT correlate more strongly to scales in CLASS®. The results indicate that the observations measured with EduCAT tap similar aspects of classroom climate quality to those measured with CLASS®, thereby confirming the new protocol's concurrent validity and instilling confidence in its use.

KEYWORDS

social-emotional climate, K-3 classrooms, interactions, assessment, EduCAT, CLASS, concurrent validity

Introduction

A positive classroom social–emotional climate (CSEC) is key to enhancing students' wellbeing and learning ability (Greenberg, 2023). As a measure of the social–emotional climate, a CSEC instrument can provide valuable information for practices and research, enabling an accurate assessment of the quality of the social–emotional climate. This information can be used to improve the classroom's educational climate and demonstrate outcomes of interventions, provided that these measures are reliable, valid, and sensitive to change (Allodi and Ringer, 2024). The social–emotional climate in the classroom is largely shaped by the interaction and relationships between the teacher and the students (CASEL,

2020). Key aspects of the CSEC include instructional, behavioral, and emotional support (CASEL, 2020). These elements, relationships, behaviors, and rules can also be the focus of tier-1 interventions in a classroom setting. Numerous programs and interventions have been developed in this area, as outlined in the CASEL framework for advancing social and emotional learning in education (CASEL, 2020). Such programs could be advantageous in our context, where evaluations report higher rates of negative classroom climate than in other countries. Analyses from the Programme for International Student Assessment, PISA 2022 (OECD, 2023) present the school disciplinary climate using an index based on 15-year-old students' perceptions of their learning environment during mathematics lessons. Sweden reports the lowest index (−0.32) among the Organization for Economic Co-operation and Development (OECD) countries in Europe (OECD, 2023, p. 353). However, the formulation of the items in the Swedish student questionnaire was changed, suggesting that these results may not be valid, or may not be as extreme as they appear, or may not be directly comparable with those of other countries (Vlachos, 2024).

Nevertheless, there are other indications of disruptions in classrooms in our context: less than 40% of students ($N = 147,000$) in compulsory schools report that they cannot work well (*Swedish = studiero*) always or often during lessons in a recent official evaluation (Swedish School Inspectorate, 2024). Interventions to improve the learning environment based on CSEC programs are not systematically employed in Sweden, despite the fact that a randomized controlled trial (RCT) of the preschool Promoting Alternative THinking Strategies (PATHS) curriculum has been conducted (Eninger et al., 2021), positive behaviour, supportive learning environment and interaction in school (PALS) are ongoing (Fohlin et al., 2023; Berg et al., 2024). There is also a lack of validated CSEC instruments, which, having a theoretical ground, can provide more than a mere score as evaluation and input for reflections on possible improvement goals. The Educational Climate Assessment Tool (EduCAT) (Allodi and Ringer, 2024) is an observational instrument with contents that are both based on theories and tested empirically, which we hypothesize captures significant and visible aspects of the CSEC.

Students' relationships with their teacher and one another and classroom dynamics can have a profound impact, either positively or negatively, on the student's learning and development (CASEL, 2020; Kiuru et al., 2015; Pakarinen et al., 2014). A negative classroom learning environment is associated with children having more emotional and behavioral problems (Milkie and Warner, 2011). According to a report by the Swedish government (SOU, 2021: 11), environments characterized by positive relationships are protective of mental health. This fact underscores the critical role of positive relationships and classroom dynamics in fostering a conducive learning environment.

Significant differences can exist in the quality of the social-emotional climate in educational settings (SOU, 2021). However, changing these conditions can be difficult due to a lack of knowledge about the essential elements that contribute to shaping the classroom climate day-by-day, the mechanisms that may produce changes, and the lack of readily available and valid tools that permit evaluating its quality. Several interventions and programs have been developed to improve the classroom's social-emotional climate (i.e., CASEL, 2020). There is still a need to evaluate the outcomes of such interventions and programs effectively: asking teachers and students about their opinions is an option, but there is also a need both in research and in

the practice of evaluations based on external observations. The validated EduCAT observation protocol could be a tool for educational practice as an assessment instrument and research, which may facilitate the planning of improvements in the CSEC (Allodi and Ringer, 2024).

Aim and rationale

The primary aim of this study was to evaluate the concurrent validity of the EduCAT protocol. This type of validity testing is conducted with a new instrument compared to an existing one to examine if the measurements produce similar results. EduCAT involves examining correlations between the scores from the EduCAT scales and scores from the Classroom Assessment Scoring System (CLASS®, Hamre et al., 2007) administered concurrently. CLASS is chosen as a comparison instrument since it is a well-established and valid instrument widely used in research and practice to measure the quality of the educational climate in the classroom, which involves relationships, management, and instructional aspects. CLASS® has been developed in the USA but has proven suitable for observing classrooms even in other contexts (Castro et al., 2017; Chan et al., 2023; Gamlem, 2019; Karlsen et al., 2024; Salminen et al., 2012; Siekkinen et al., 2013; Silinskas et al., 2017; Virtanen et al., 2018). The CLASS® instrument, however, is not freely available open access; to be a qualified rater requires comprehensive training, which is offered only in the USA, to our knowledge; after the first training, the authorized raters have to update their qualification each year. These characteristics and the costs related to the training have made the CLASS® tool easier to apply in research than in school practice, where the resources are limited. The newly developed EduCAT instrument will be open access, even if raters' training is necessary. If validated, EduCAT could become a complement to CLASS® in research and evaluations in our context and an alternative, for example, to measure outcomes of classroom climate interventions in the school practice. The successful validation of EduCAT could pave the way for more effective interventions and strategies to enhance students' learning and development.

To test the concurrent validity of EduCAT, we will

- Analyze the contents of the two instruments and show how they assess similar constructs;
- Examine the psychometric properties of the two instruments reporting descriptive statistics and reliability for the scales;
- Conduct pairwise correlations between the scales;
- Conduct canonical correlations analysis (CCA) of the scales from EduCAT and CLASS. CCA is chosen since it can be employed to investigate to which extent one set of variables (EduCAT) can be predicted by another set of variables (CLASS) (Thompson, 1984).

Previous research and development of EduCAT

The EduCAT protocol (Allodi and Ringer, 2024) has been developed in Sweden as an observer-rated classroom social-emotional climate

index. EduCAT has been developed based on a framework incorporating theories of learning environments, motivation, universal values, student evaluations (Allodi, 2002, 2010a), and teacher interviews (Allodi and Ringer, 2024). This CSEC measure is grounded on a conceptual model of classroom climate (Allodi, 2010b) that applies qualitative analysis of students' narratives and evaluations regarding their experiences at school (Allodi, 2002). The empirical lived experiences of what students value in the classrooms were linked to the framework of the psychosocial environment (Moos, 1979), the universal theory of motivational values (Schwartz, 1992), and classroom climate research (Fraser, 1986, 1991). The model was operationalized as a student questionnaire (Allodi, 2007), with 40 items related to 10 correlated factors (creativity, stimulation, achievement, self-efficacy, safety, control, helpfulness, participation, responsibility, and influence). The structural model showed a good fit in previous data collections (Allodi, 2007, 2010b). A 50-item teacher self-assessment questionnaire used the same 10 factors, which showed good ecological validity in small trials (Wikman, 2023). Following these studies, an observation protocol with 5 factors and 20 items in a 4-grade Likert scale, based on the same framework and 2 waves of practitioners' feedback, has been devised and tested in various schools and grades (Allodi and Ringer, 2022, 2023, 2024).

Method

Participants

A convenience sample of four schools was recruited: three municipal and one independent. Eligibility was grade kindergarten to third grade (K-3). In total, 30 teachers from grades K-3 consented to participate. The sample included 1 kindergarten, 10 first-grade, 11 second-grade, and 8 third-grade classes.

Procedures

A trained CLASS and EduCAT observer collected data during regular school hours. After attending a workshop on the CLASS system, the observer passed a reliability test within a point of the master-coded tapes on 80% of the scores across five video films. During the certification renewal period, the researcher regularly passed this test. The observer (first author), in addition to training in CLASS and EduCAT, has 17 years of experience as a teacher and special educator. The CSEC was assessed with the two instruments close in time. According to the CLASS protocol, the observer thrice assessed the CSEC for 20 min. In total, each class was observed for 60 min. Following this, the same 60-min lesson was assessed using the EduCAT protocol. The observer moved freely in the classroom during the observation without interacting with students or teachers. The anonymized numerical data from the observations were stored in secure digital repositories and are only accessible to authorized researchers.

Educational climate assessment tool (EduCAT)

The previous version of EduCAT (Allodi and Ringer, 2022) consisted of 50 items corresponding to 10 dimensions, rated on a 7-point Likert scale. The revised version of EduCAT has five scales,

each with four indicators; the total number of items is 20, rated on a 4-point Likert scale. The reduction of the number of scales and items was directed by the goal to increase the ecological validity and perceived usability of the instrument by practitioners; the reduction of the rating scale range was also directed by the consideration to be easier to achieve inter-rater agreement on the items' scoring. The formulations of the 4-point Likert scale are as follows: 1. Does not occur during any activity; 2. Occurs to a limited extent in any activity; 3. Occurs in many activities; 4. Occurs in all activities. Each scale ranges from 4 to 16 points, and the total EduCAT protocol scores range from 20 to 80.

Contents of EduCAT

The 20 items represent 5 dimensions of CSEC (stimulation, improvement, structure, relationship, and influence). The *stimulation* scale targets active participation, expressions of enthusiasm, playful and creative work, and personal experiences. The *improvement* scale includes the level of encouragement the teacher provides, the feedback given, and whether the students are encouraged to make an effort despite facing complex tasks. The *structure* scale targets the level of clear rules and structure in the classroom and the order during the activities. The scale of *relationship* measures how students help each other emotionally or with tasks, share freely in the classroom, and show openness to one another. The scale of *influence* encompasses students' opportunities to express suggestions, choose between alternatives, and be assigned also adequate responsibilities. The scales are supposed to capture distinct but correlated aspects of the classroom's social emotional, and educational climate.

Construct validity of EduCAT

The construct validity of EduCAT was tested in another sample ($N = 53$ K-3 classrooms). The internal reliability of the five scales ranged from 0.90 to 0.94, as measured by Cronbach's α . The structure of EduCAT was obtained from a principal component analysis that identified five factors with an eigenvalue above 1 that explained 84% of the variance. The four items in each dimension had high factor loadings (from 0.65 to 0.90) on five factors corresponding to the hypothesized structure (Allodi and Ringer, 2024).

Classroom assessment scoring system

CLASS (Hamre et al., 2007) assesses the CSEC by emphasizing interactions between teachers and students as the principal mechanism of student learning and development (Hamre and Pianta, 2005). This instrument contains three primary domains and 10 dimensions of positive student outcomes. An assessment of each dimension is conducted on a 7-point Likert scale, with 1 and 2 being low; 3–5 being mid-range; and 6 and 7 being high. For each domain, an average score is calculated. A negative climate score of eight is subtracted from the average score of NC to obtain a reversed score. The maximum score is 70 (Pianta et al., 2008).

Contents of CLASS

The *emotional support domain* comprises four dimensions: positive climate (PC), negative climate (NC), teacher sensitivity (TS), and regard for students' perspectives (RSP). These dimensions assess the amount and quality of positive and negative interactions among students, the teacher's awareness of the students' emotional and academic needs, and the teacher's focus on students' autonomy.

The *classroom management domain* includes three dimensions: behavior management (BM), productivity (PD), and instructional learning formats (ILF). This domain assesses a teacher's ability to manage student behavior proactively, utilize time effectively, and retain students' interest.

The *instructional support domain* comprises three dimensions: concept development (CD), quality of feedback (QF), and language modeling (LM). A teacher's proficiency in the instructional domain can be evaluated based on their ability to use thinking skills, provide feedback effectively, and promote the development of language skills (Pianta et al., 2008).

Observation procedures of CLASS

Observations can last from one to six cycles, with each observation period lasting 20 min (Pianta et al., 2008). In the present study, a researcher trained to rate classrooms observed each classroom for three cycles. The present study focuses on the K-3 version. Based on past research with CLASS, the intraclass correlation coefficients (ICCs) for emotional support were 0.77, classroom organization was 0.82, and instructional support was 0.73. The ICCs for the dimensions ranged from 0.64 to 0.78, except for ICCs of 0.50 for negative climate (Allen et al., 2013).

Statistical analysis

Analyses were conducted using Statistical Package for the Social Sciences (SPSS) version 29.0 software (SPSS, 2023). At the observed level, normality (item and scale), item-total correlations, and scale reliability were examined (Cronbach's α and Omega). Canonical correlation analysis (CCA) (Boedeker and Henson, 2020; Meyers et al., 2017) was used to explore the multivariate relationship between two sets of variables: the 10 CLASS scales and the five EduCAT scales (Table 1).

Results

Descriptive statistics

The descriptive statistics of the instruments with the mean sum scores and standard deviations are reported in Table 1. Due to zero variance, the scale negative climate (reversed) from CLASS is not reported and will not be used in the following analyses. Internal consistency for EduCAT and CLASS scales is reported in Table 1. The scale with the highest scores in EduCAT is structure, and the lowest is improvement. The other scales show scores above 2. The variance and standard deviation do not differ much. The internal consistency of the

TABLE 1 Descriptive statistics for EduCAT and CLASS: item mean, standard deviation, and scale internal consistency (Cronbach alpha and omega).

Domains	Mean	SD	Internal consistency alpha	Internal consistency omega
EduCAT				
STI (items 1, 2, 4)	2.66	0.57	0.54	0.59
IMP	1.85	0.65	0.64	0.63
STR	3.28	0.88	0.92	0.94
REL	2.37	0.73	0.80	0.81
INF	2.25	0.67	0.68	0.71
CLASS				
PC	5.41	1.39	0.97	0.97
TS	5.34	1.46	0.94	0.94
RSP	4.37	1.14	0.96	0.96
BM	5.65	1.68	0.98	0.98
PD	5.81	1.50	0.94	0.94
ILF	4.82	1.05	0.94	0.94
CD	2.21	0.86	0.82	0.83
QF	2.81	0.93	0.75	0.81
LM	2.34	0.90	0.83	0.85

EduCAT: STI, stimulation; IMP, improvement; STR, structure; REL, relationship; INF, influence.

CLASS: PC, positive climate; TS, teacher sensitivity; RSP, regard for students perspectives; BM, behavior management; PD, productivity; ILF, instructional learning formats; CD, concept development; QF, quality of feedback; LM, language modeling. The Negative Climate scale was not included due to zero variance.

scales is good for structure and relationship but lower for improvement and influence. Stimulation shows poor internal consistency with Cronbach's α and ω , where item 3 (the work is done playfully and creatively) was not positively correlated to the other items.

The scales with the higher scores from CLASS are behavior management and productivity, but even a positive climate shows high scores. The scales on the instructional support domain, concept, feedback, and language show the lowest scores. All the scales in CLASS show good internal consistency.

Predicted relationships and correlations

The comparison instrument CLASS was selected based on its reliability and validity reported for age groups appropriate to EduCAT and its similar conceptual constructs. Both instruments were developed as dimension-specific instruments for assessing CSEC. The general hypothesis was that if a classroom got high scores with CLASS, the EduCAT scores would also be high, and vice versa. More specifically, the following expected relationships were derived based on a comparison of the contents of the CLASS questionnaire and the EduCAT scales. Positive correlations were expected between CLASS and EduCAT dimensions, except for concept development and language modeling as these constructs are not covered in EduCAT. The CLASS dimensions of positive and teacher sensitivity

TABLE 2 Spearman’s correlation coefficients relate the EduCAT to the subscales of the CLASS instrument.

Domains	Stimulation	Improvement	Structure	Relationship	Influence
PC	0.32	0.58	0.73	0.34	0.41
TS	0.46	0.69	0.78	0.31	0.43
RSP	0.38	0.34	0.27	0.45	0.94
BM	0.37	0.47	0.87	0.43	0.44
PD	0.39	0.52	0.87	0.24	0.41
ILF	0.89	0.63	0.58	0.32	0.40
CD	0.59	0.73	0.30	0.35	0.37
QF	0.60	0.96	0.57	0.33	0.43
LM	0.51	0.71	0.43	0.21	0.38

PC, positive climate; TS, teacher sensitivity; RSP, regard for students perspectives; BM, behavior management; PD, productivity; ILF, instructional learning formats; CD, concept development; QF, quality of feedback; LM, language modeling.

All the correlations between the scales in CLASS and EduCAT are positive.

The values above 0.70 are in bold and indicate a strong correlation.

were expected to correlate positively with the EduCAT dimension of relationship. Behavior management and productivity were expected to correlate strongly with structure; instructional learning format was expected to correlate strongly with stimulation; quality of feedback was expected to correlate with improvement; and regard for students’ perspective was expected to correlate strongly with influence.

The means of the scores for each scale in both instruments were standardized, and the correlations between these values were calculated. Table 2 presents Spearman’s correlation coefficients between standardized mean scores on EduCAT scales and CLASS subscales. The hypothesized strong correlations were confirmed in the majority of cases: for stimulation with instructional learning formats, improvement with feedback (and additionally with other scales—concept development and language modeling), and influence with regard to students. Still, the correlations between positive climate and teacher sensitivity were only moderate for relationships. The improvement scale showed strong correlations even with other scales in CLASS, particularly all four scales related to instruction and teaching; the structure scale also taps into the CLASS scales about positive climate and teacher sensitivity. However, relationship shows positive but only moderate correlations to scales in CLASS. It could be viewed as a scale that targets behaviors somewhat different from those targeted in CLASS.

An exploration with canonical correlation analysis

The multivariate relationship between CLASS and EduCAT can be explored further by CCA. The results of the canonical correlation analysis between the standardized values of the average scores of the nine scales of CLASS and the standardized values of the average scores of the five scales of EduCAT showed five canonical correlations, of which four were significant canonical correlations and they are $\rho_1 = 0.99$ ($p < 0.001$); $\rho_2 = 0.98$ ($p < 0.001$), $\rho_3 = 0.91$ ($p < 0.001$), and $\rho_4 = 0.83$ ($p < 0.003$).

The canonical loadings of the standardized variables from CLASS and EduCAT (1–4) are reported in Table 3 (CLASS) and 4 (EduCAT) and in Figure 1 report the canonical loadings of the

TABLE 3 Canonical loadings to the four canonical variates from the CLASS scores.

Variable	V1	V2	V3	V4
PC	0.27	0.65	0.44	−0.14
TS	0.25	0.79	0.33	−0.16
RSP	0.99	0.00	0.05	−0.01
BM	0.19	0.75	0.61	−0.04
PD	0.20	0.76	0.59	−0.07
ILF	0.41	0.80	0.11	0.39
CD	0.29	0.64	−0.38	0.05
QF	0.31	0.83	−0.33	−0.30
LM	0.31	0.61	−0.25	−0.00

PC, positive climate; TS, teacher sensitivity; RSP, regard for students perspectives; BM, behavior management; PD, productivity; ILF, instructional learning formats; CD, concept development; QF, quality of feedback; LM, language modeling.

standardized variables from CLASS and EduCAT. The first variate pair contributes most *regard to students’ perspectives* and *instructional learning formats* from CLASS; *Influence* contributes most from EduCAT, and moderately from *relationship*. The interpretation of these variates is that they represent the extent to which the classroom is child-friendly, where children are listened to by the teacher and are kind to each other. To the second variate pair, the contribution from CLASS comes from all the other eight scales, except with *regard to student perspectives*. From EduCAT, the contribution to canonical variate 2 comes from *improvement, structure, and stimulation*. The interpretation of these variates is that they represent *the overall climate conducive to learning*. To the third variate pair, the contribution from CLASS comes mainly from *behavior management* and *production* and EduCAT only from *structure*. These can be interpreted as *classroom management and routines*. To the fourth variate pair the contribution is from *CLASS instructional learning formats* and EduCAT *stimulation*. It can be interpreted as *engaging instruction and activities* (Table 4).

Table 5 presents the proportion of standardized variance of each set of original variables explained by canonical variates. The cumulative variance explained by four canonical variates for set 1 is

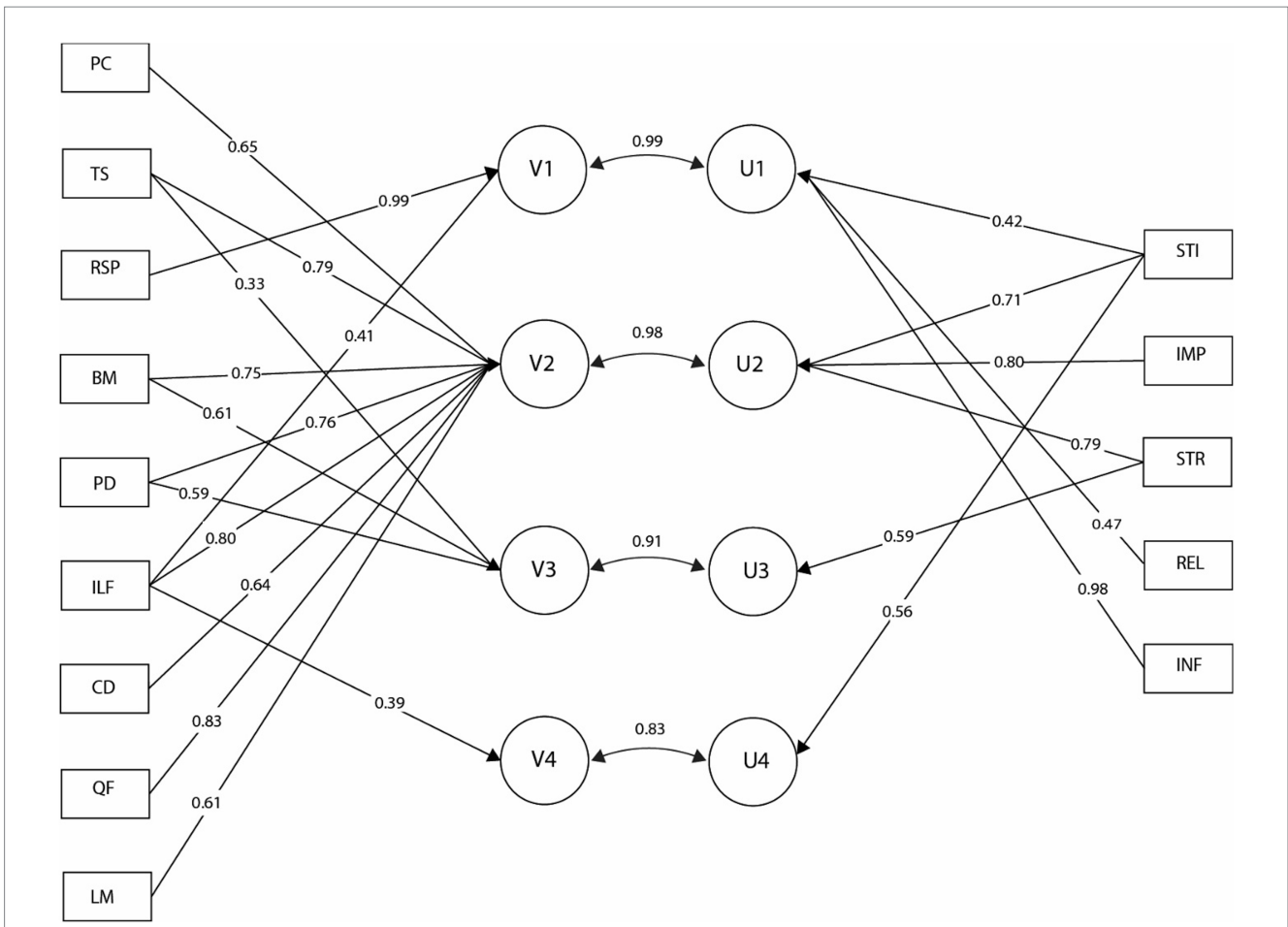


FIGURE 1 Canonical correlations analysis (CCA) of the two sets of scales from EduCAT and CLASS. Four canonical variate pairs V1-U1, V2-U2, V3-U3, V4-U4 are represented as circles; The canonical correlations between V and U are the numbers above bowed lines with double arrows. The variables in set 1 are represented by rectangles on the left side and they are: PC, positive climate; NC, negative climate; TS, teacher sensitivity; RSP, regard for students perspectives; BM, behavior management; PD, productivity; ILF, instructional learning formats; CD, concept development; QF, quality of feedback; LM, language modeling. The canonical loadings to the four canonical variates from the CLASS scores are represented by arrows from set 1 to V1-V4. The variables in set 2 are represented by rectangles on the right side and are: STI, stimulation; IMP, improvement; STR, structure; REL, relationship; INF, influence. The canonical loadings to the canonical variates from EduCAT scores are represented by arrows from set 2 to U 1-4.

TABLE 4 Canonical loadings to the canonical variates from EduCAT scores.

Variable	U1	U2	U3	U4
STI	0.42	0.71	-0.05	0.56
IMP	0.18	0.80	-0.33	-0.46
STR	0.15	0.79	0.59	-0.07
REL	0.47	0.27	0.12	0.06
INF	0.98	-0.04	0.10	-0.08

STI, stimulation; IMP, improvement; STR, structure; REL, relationship; INF, influence.

TABLE 5 Redundancy analysis: proportion of variance explained by canonical variates (1-5).

Variable	Set 1 by Self	Set 1 by Set 2	Set 2 by Self	Set 2 by Set 1
1	0.18	0.18	0.28	0.27
2	0.48	0.45	0.37	0.35
3	0.15	0.13	0.10	0.08
4	0.03	0.02	0.11	0.07
5	0.01	0.00	0.15	0.03

Set 1 = CLASS; Set 2 = EduCAT.

86%; the cumulative variance explained by four canonical variates for set 2 is 85.6%.

Discussion

In this study, the educational climate was observed with two instruments by a researcher with broad experience in teaching pupils at

that age; a limitation is that the same observer conducted the observations. The availability of two observers who could observe the classroom simultaneously and rate independently on CLASS and EduCAT would have strengthened the validity of the data collection. The observer was aware of the importance of scoring's impartiality and tried to avoid bias between the ratings of the two instruments. The items in EduCAT and CLASS have, to a great extent, different formulations, which may make it

TABLE 6 Significance tests of canonical correlations.

Variable	Correlation	Eigenvalue	Wilks Statistic	F	p values
1	0.99	33.57	0.00	12.91	<0.00
2	0.98	19.80	0.00	8.71	<0.00
3	0.91	4.79	0.04	4.94	<0.00
4	0.83	2.17	0.25	3.16	<0.00
5	0.45	0.26	0.80	1.03	<0.43

H0 for the Wilks test is zero canonical correlations in the current and following rows.

possible to perform the second scoring without being influenced too heavily by the first one. In future studies, one aim would be to conduct independent observations. Another limitation is that the sample was small. In CCA, larger samples are recommended to achieve robust results that are stable at replications. However, for exploratory studies, the requirements of CCA can be relaxed (Hair et al., 2010, p. 243). We consider the CCA results as a promising first exploration of the multivariate relationships between the observation instrument CLASS and EduCAT. We also consider that further analyses with a larger sample should be conducted (Table 6).

The descriptive statistics show that the CLASS scales have high internal reliability, while the EduCAT scales have lower internal reliability in these observations. Since previous observations in K-3 classrooms reported higher internal reliability (Allodi and Ringer, 2024), we considered the item means for five scales still acceptable in these exploratory analyses. Reliability values of 0.60 to 0.70 are deemed the lower limit of acceptability (Hair et al., 2010, p. 92). The reliability of stimulation has to be looked at, particularly the content of item 3 (the work is done playfully and creatively), as it seemed to correlate poorly with the other items in the scale in this set of observations.

The pairwise correlations and the canonical correlation analysis of the standardized scores (item means) of the scales from CLASS and EduCAT show that the scales are correlated and tap similarly various aspects of the events, activities, and relationships that take place in the classroom, observed by an external observer, trained to detect relevant behaviors and events.

The canonical correlation analysis shows that the first canonical variates are related to a *child-friendly climate*, where the teacher pays attention to the children's views. The *overall climate conducive to learning* was the interpretation of the second variates. Interestingly, the third function that seems to represent *behavior management and routines* accounts for only a minor variance (almost 10%), even if these classroom climate aspects are usually presented as central in various contexts, such as teacher training or professional development. They indeed play a distinct role, but this analysis suggests instead that the social and emotional aspects of the educational environment that can be associated with a child-friendly climate seem more meaningful in differentiating learning environments in elementary classrooms in our context.

These results indicate that the newly developed EduCAT observation tool has concurrent validity with CLASS, a well-established measure of the quality of the learning environment. The EduCAT seems to be a valid and promising tool that can be tested in field trials and in collaboration with practitioners, even if the reliability of the scales could be improved. The following steps in the practice field would include the development of a manual on the theoretical framework and the contents of EduCAT, which would include strategies to improve the classroom climate. In future research, the

training of observers could be developed, and inter-rater agreement and social validity investigations could be conducted to prepare for trials. The sensitivity of EduCAT to interventions could be tested in field trials, along with its predictive validity of student outcomes such as wellbeing and achievement.

The contents of EduCAT have been developed based on theories, empirical data analysis, and practitioners' feedback. They are probably well-suited to the Swedish educational context, curriculum, and standard educational practices. Trying EduCAT in other contexts would be possible, but it is essential to know that some adaptations may be needed.

In comparison to other educational contexts, the learning environment and social climate in Swedish classrooms show significant variation; a low disciplinary climate is reported to a greater extent than in similar countries with high educational standards and resources, based on international assessments (OECD, 2023), even if the validity of the comparison can be questioned due to changes in the questionnaire items (Vlachos, 2024). School or classroom disciplinary climate is another way, perhaps more negatively connotated, to define the characteristics of the learning environment, which are also targeted by the observation tools that we have analyzed in this paper (they are targeted, for example, in behavior management and structure). However, the quality of the learning environment is more than just discipline, considering that the term *discipline* may be associated with punishment. In our view, the quality of the learning environment is a multifaceted construct, including instruction, learning, management, and social and emotional aspects. Nevertheless, the low scores on school disciplinary climate in our context should be investigated further, as they may raise concerns for policymakers and stakeholders, even if the issues are not as extreme as they appear from the PISA comparison (Swedish School Inspectorate, 2024). A scientifically valid assessment of the quality of learning environments, one that is socially valid and accepted by the teachers, could be effectively used in interventions that improve the classroom climate within the Swedish educational system.

Data availability statement

The datasets presented in this article are not readily available because this study's ethical review does not allow the primary data to be in a public repository. Individual data requests for de-identified (IPD) will be reviewed for qualified researchers (e.g., Ph.D.) who obtain ethical permission under Swedish ethical regulations/laws for secondary data analysis for purposes such as meta-analysis or confirmation of published study results. Requests to access the datasets should be directed to carina.wikman@speccped.su.se.

Author contributions

CW: Conceptualization, Investigation, Data curation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing. MA: Funding acquisition, Conceptualization, Formal analysis, Writing - original draft, Writing - review & editing. NR: Conceptualization, Writing - review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by Swedish Research Council 2018–04012.

References

- Allodi, M. W. (2002). Children's experiences of school: Narratives of Swedish children with and without learning difficulties. *Scand. J. Educ. Res.* 46, 181–205. doi: 10.1080/00313830220142191
- Allodi, M. W. (2007). Assessing the quality of learning environments in Swedish schools: Development and analysis of a theory-based instrument. *Learn. Environ. Res.* 10, 157–175. doi: 10.1007/s10984-007-9029-9
- Allodi, M. W. (2010a). The meaning of social climate of learning environments: Some reasons why we do not care enough about it. *Learn. Environ. Res.* 13, 89–104. doi: 10.1007/s10984-010-9072-9
- Allodi, M. W. (2010b). Goals and values in school: A model developed for describing, evaluating, and changing the social climate of learning environments. *Soc. Psychol. Educ.* 13, 207–235. doi: 10.1007/s11218-009-9110-6
- Allodi, M. W., and Ringer, N. (2022). Students' well-being and perceptions of social climate measured with GAVIS- a theory -based instrument. Presented at the European Conference on Positive Psychology (ECP2022), 29 June- 2 July, 2022, Reykjavik, Iceland.
- Allodi, M. W., and Ringer, N. (2024). GAVIS Observation Protocol 2024 (Version 1). *Stockholm University*. doi: 10.17045/sthlmuni.25204799.v1
- Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., and Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: predicting student achievement with the classroom assessment scoring system--secondary. *Sch. Psychol. Rev.* 42, 76–98. doi: 10.1080/02796015.2013.12087492
- Berg, K. E. C., Lundgren, J. S., and Bernhardsson, S. (2024). Facilitators and barriers for implementing the PALS school-wide positive behavior support model in a Swedish municipality: a focus group study. *Scand. J. Psychol.* 65, 919–936. doi: 10.1111/sjop.13045
- Boedeker, P., and Henson, R. (2020). "Canonical correlation analysis" in eds. P. Atkinson, S. Delamont, A. Cernat, J. W. Sakshaug and R. A. Williams (SAGE Research Methods Foundations).
- CASEL. Collaborative for Academic, Social, and Emotional Learning. CASEL's SEL Framework. (2020). Available at: <https://casel.org/casel-sel-framework-11-2020/>
- Castro, S., Granlund, M., and Almqvist, L. (2017). The relationship between classroom quality-related variables and engagement levels in Swedish preschool classrooms: a longitudinal study. *Eur. Early Child. Educ. Res. J.* 25, 122–135. doi: 10.1080/1350293X.2015.1102413
- Chan, S. W., Pöysä, S., Lerkkanen, M.-K., and Pakarinen, E. (2023). Teachers' occupational well-being in relation to teacher-student interactions at the lower secondary school level. *Scand. J. Educ. Res.* 68, 1137–1154. doi: 10.1080/00313831.2023.2204114
- Eninger, L., Ferrer-Wreder, L., Eichas, K., Olsson, T. M., Hau, H. G., Allodi, M. W., et al. (2021). A Cluster Randomized Trial of Promoting Alternative Thinking Strategies (PATHS[®]) With Swedish Preschool Children. *Front Psychol.* 12, 695288. doi: 10.3389/fpsyg.2021.695288
- Fraser, B. J. (1986). *Classroom environment*. London: Croom Helm.
- Fohlin, L., Westling Allodi, M., Sedem, M., and Karlberg, M. (2023). School staff's perceptions of implementing the Inclusive Behavioral Support in Schools framework in Swedish schools. *Scand. J. Educ. Res.* 1–17. doi: 10.1080/00313831.2023.2287437
- Fraser, B. J. (1991). "Two decades of classroom environment research" in *Educational environments. Evaluation, antecedents, and consequences*. eds. B. J. Fraser and H. J. Walberg (Oxford: Pergamon Press), 3–27.
- Gamlem, S. M. (2019). Mapping teaching through interactions and pupils' learning in mathematics. *SAGE Open* 9, 1–13. doi: 10.1177/2158244019861485
- Greenberg, M. T. (2023). Evidence for social and emotional learning in schools. *Learning Policy Inst.* doi: 10.54300/928.269
- Hair, J. F., Black, W., Babin, B., and Anderson, R. (2010). *Multivariate data analysis: A global perspective*. 7th Edn. Upper Saddle River, NJ: Pearson Education.
- Hamre, B. K., and Pianta, R. C. (2005). Can instructional and emotional support in the first grade classroom make a difference for children at risk of school failure? *Child Dev.* 76, 949–967. doi: 10.1111/j.1467-8624.2005.00889.x
- Hamre, B. K., Pianta, R. C., Mashburn, A. J., and Downer, J. T. (2007). Building a science of classrooms: application of the CLASS framework in over 4,000 US early childhood and elementary classrooms. *Found. Child. Dev.* 30.
- Karlsen, L., Rydland, V., Solheim Buøen, E., Vandell, D. L., and Lekhal, R. (2024). The factor structures of the classroom assessment scoring system pre-K and mature play observation tool in multi-ethnic Norwegian early childhood centers. *J. Early Child. Res.* 22, 106–121. doi: 10.1177/1476718X231195708
- Kiuru, N., Aunola, K., Lerkkanen, M., Pakarinen, E., Poskiparta, E., Ahonen, T., et al. (2015). Positive teacher and peer relations combine to predict primary school students' academic skill development. *Dev. Psychol.* 51, 434–446. doi: 10.1037/a0038911
- Meyers, L., Gamst, G., and Guarino, A. (2017). "Canonical correlation analysis using IBM SPSS" in *Applied multivariate research: Design and interpretation* (SAGE Publications, Inc).
- Milkie, M. A., and Warner, C. H. (2011). Classroom learning environments and the mental health of first grade children. *J. Health Soc. Behav.* 52, 4–22. doi: 10.1177/0022146510394952
- Moos, R. H. (1979). *Evaluating educational environment*. San Francisco: Jossey-Bass.
- OECD (2023). PISA 2022 results (volume II) learning during – And from–disruption. Available at: https://www.oecd-ilibrary.org/education/pisa-2022-results-volume-ii_a97db61c-en
- Pakarinen, E., Aunola, K., Kiuru, N., Lerkkanen, M., Poikkeus, A., Siekkinen, M., et al. (2014). The cross-lagged associations between classroom interactions and children's achievement behaviors. *Contemp. Educ. Psychol.* 39, 248–261. doi: 10.1016/j.cedpsych.2014.06.001
- Pianta, R. C., La Paro, K. M., and Hamre, B. K. (2008). *The classroom assessment scoring system manual, K-3*. Baltimore, MD: Brookes.
- Salminen, J., Lerkkanen, M.-K., Poikkeus, A.-M., Pakarinen, E., Siekkinen, M., Hännikäinen, M., et al. (2012). Observed classroom quality profiles of kindergarten classrooms in Finland. *Early Educ. Dev.* 23, 654–677. doi: 10.1080/10409289.2011.574267
- Schwartz, S. H. (1992). "Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries" in *Advances in experimental psychology*. ed. M. P. Zanna (London: Academic Press), 1–65.
- Siekkinen, M., Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., Salminen, J., Poskiparta, E., et al. (2013). Social competence among 6-year-old children and classroom instructional support and teacher stress. *Early Educ. Dev.* 24, 877–897. doi: 10.1080/10409289.2013.745183
- Silinskas, G., Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., and Nurmi, J.-E. (2017). Classroom interaction and literacy activities in kindergarten: longitudinal links to grade 1 readers at risk and not at risk of reading difficulties. *Contemp. Educ. Psychol.* 51, 321–335. doi: 10.1016/j.cedpsych.2017.09.002

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- SOU. (2021). Bättre möjligheter för elever att nå kunskapskraven – aktivt stöd-och elevhälsoarbete samt stärkt utbildning för elever med intellektuell funktionsnedsättning. Slutbetänkande av utredningen om elevers möjligheter att nå kunskapskraven. Final report of the investigation into student opportunities to meet the knowledge requirements. 11. Available at: <https://www.regeringen.se/492cd0/contentassets/77f91fad3ff4e9f85bc560b1e5c34d9/battre-mojligheter-for-elever-att-na-kunskapskraven-sou-202111>
- SPSS. (2023). Statistics for windows, version 29.0. Chicago: SPSS Incorporate.
- Swedish School Inspectorate (2024). Trygghet och studiero 2023 [safety and study peace]. Available at: <https://www.skolinspektionen.se/skolenkaten>
- Thompson, B. (1984). Canonical correlation analysis [Elektronisk resurs] uses and interpretation. SAGE: Newbury Park.
- Virtanen, T. E., Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., Siekkinen, M., and Nurmi, J.-E. (2018). A validation study of classroom assessment scoring system–secondary in the Finnish school context. *J. Early Adolesc.* 38, 849–880. doi: 10.1177/0272431617699944
- Vlachos, J. (2024). Ordning eller översättningsproblem i PISA enkäter [Discipline or translation problems in PISA questionnaires]. *Ekonomistas*. Available at: <https://ekonomistas.se/2024/06/18/ordnings-eller-oversattningsproblem-i-pisa-enkater/>
- Wikman, C. (2023). *Social Climate and the Student in the Learning Environment: Advances in Assessment, Observation, and Coaching* (PhD dissertation, Department of Special Education, Stockholm University) <https://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-220641>