Education in the North

UNIVERSITY OF ABERDEEN

ARTICLE

The malleability of students' perception of key teaching factors within and across Nordic contexts: validating the 7Cs framework

Eva Lykkegaard, eval@sdu.dk
University of Southern Denmark, Denmark
https://orcid.org/0000-0002-2266-9528

Ane Qvortrup, anq@sdu.dk
University of Southern Denmark, Denmark
https://orcid.org/0000-0002-2110-0575

DOI Number: https://doi.org/10.26203/n7tv-6n78

To cite this article: Lykkegaard, E. and Qvortrup, A., (2024). The malleability of students' perception of key teaching factors within and across Nordic contexts: validating the 7Cs framework. *Education in the North*, **31**(1) pp.79-105.

The malleability of students' perception of key teaching factors within and across Nordic contexts: validating the 7Cs framework

Eva Lykkegaard, eval@sdu.dk

University of Southern Denmark, Denmark

Ane Qvortrup, anq@sdu.dk

University of Southern Denmark, Denmark

Abstract

The 7Cs framework measures seven key pedagogical factors (Care, Control, Clarify, Challenge, Captivate, Confer, Consolidate) and is one of the most popular student perception surveys to measure quality in the United States. The framework is originally said to be a valid and reliable indicator of teaching quality across time and contexts, but an increasing number of studies indicate that the factors are not necessarily stable, but malleable according to different contexts. Currently, the framework is extending its reach to Nordic contexts and it is thus important to examine whether the factors are stable across different Nordic Countries. Thus, in this paper, we validate the 7C framework for Nordic contexts and hereby examine if and how they differ across different Nordic contexts. We start by presenting our constructivist stance and the way this stance guides us in the subsequent analysis. The analysis is based on 7C surveys collected in 2020/2021 in Denmark, Norway, Sweden, Finland, and Iceland (N=2,265 upper secondary students). By following well-tested and recognized guidelines and steps for validating educational environment instruments we show that students' perceptions of the 7C teaching characteristics differ across the Nordic contexts. One factor, the Control factor, is consistent both across the different analyses in this article and previous studies, but all the other factors vary much across contexts. The article thus emphasizes that we must be careful about transferring frameworks developed in certain contexts and using them as indicators of quality in other contexts. In addition to contributing to knowledge about the use of the 7C framework in Nordic contexts, the paper in general contributes to the discussion of how we can and should discuss and examine quality in teaching.

Keywords: 7Cs framework, validation, Nordic contexts, pedagogical factors, malleability

Introduction

The 7Cs framework, which was developed by Ferguson and colleagues, is among the most popular instruments to measure quality in the United States (Rowley et al., 2019; Wallace et al., 2016). The instrument is designed to measure seven key pedagogical factors (Care, Control, Clarify, Challenge, Captivate, Confer, Consolidate), which are categorized into three domains: Personal support (Care and Confer), Curricular support (Captivate, Clarify, and Consolidate) and Academic pressure (Challenge and Control), as illustrated in Appendix 1. According to Ferguson and Danielson (2015), the pedagogical factors represent "multiple aspects of teaching that together predict student engagement and learning" (Ferguson and Danielson, 2015, p.98) that are stable across U.S. contexts (Ferguson, 2012) and time (Rowley et al., 2019). Thus, the factors are suggested to be used as general and "valid and reliable indicators of teaching quality" (ibid., p.101). Although several studies from researchers other than Ferguson, Danielson, and Rowling confirm the proposed stability of the 7C factors across different contexts in the United States (Bradshaw, 2017; Chaplin et al., 2014; Cohen et al., 2018; Crow, 2011; Hiver et al., 2021; Phillips et al., 2021), others do not. Wallace et al. (2016) and Kuhfeld (2016, 2017) find that a two-dimensional structure with classroom management (Control) and support (Care, Clarify, Challenge, Captivate, Confer, and Consolidate) best fit the items of the 7Cs Framework. Schweig (2014) suggests five factors at the within-school level and two factors at the between-school level. Polikoff (2015) examines the stability over time. His results also challenge the idea of stability, in that between-year stability is lower than within-year stability.

These years, interest in the 7C framework is spreading beyond United States contexts, and in light of the challenged stability within the United States alone, it is important to examine the framework outside the United States before it is simply put into use. In this article, as part of the Nordic Center of Excellence 'Quality in Nordic Teaching (QUINT)' funded by NordForsk, we examine the framework across Nordic contexts. The project formulates its vision as investigating teaching quality in the Nordic countries and intends to validate the 7C framework for comparability and cultural biases across the participating countries:

> "A validated and well-designed survey on student perceptions of teaching quality and teaching and learning processes […] will be distributed […] and validated for comparability and cultural biases across the participating countries" (QUINT-application).

Data on the 7Cs were collected among students from all five Nordic countries: Denmark, Sweden, Norway, Iceland, and Finland. We use these data to investigate whether the 7C framework is a reliable measure or is malleable across Nordic contexts. Our research question is:

How do students' perceptions of seven key teaching characteristics: Care, Confer, Captivate, Clarify, Consolidate, Challenge, and Control (The 7Cs) differ across different Nordic contexts?

Besides contributing to the Nordic validation of the 7C framework, we in the paper contribute to the field of research on teaching quality by examining and discussing how to identify relevant pedagogical factors, which are described as necessary for progression in the field (see below).

Theoretical framework

As described above, the 7C instrument is designed to measure pedagogical factors. Pedagogical factors constitute so-called changeable factors that contrasts to non-changeable or given, endogenous, "prior" conditions or factors such as student demographics, students' previous educational results, admission criteria, housing situation, etc. (Scheerens and Marks, 2017). The idea that factors are changeable refers to the fact that one can actually intervene on these factors (Yik et al., 2022). Changeable factors include educational programs, policies, and institutional factors, such as class size, student–staff ratios, indoor environment, length of the school day, and pedagogical factors, that is all types of teaching practices and activities (Scheerens, 2017). There is widespread agreement that changeable factors are important to student achievement, and that the strongest basis for understanding and strengthening the quality of teaching is established when changeable factors are given particular awareness (Hanushek, 2011; Scheerens, 2017). However, it has been difficult in reliable ways to identify specific factors with a reasonable effect size (Hanushek, 2011; Rivkin et al., 2005; Rockoff et al., 2011; Scheerens, 2017), which is suggested to be a crucial reason for the lack of progress in the research area (Archer et al., 2014; Borman et al., 2003; Detterman, 2016; Ferguson and Hirsch, 2014; Johnson, 2006; Rivkin et al., 2005; Rockoff et al., 2011; Scheerens, 2014, 2017; Scheerens and Marks, 2017; Timmermans, 2012). It is argued that a core reason for the difficulty is that the changeable factors are malleable in the sense that they shape themselves according to the context (de Boer et al., 2014; Edmonds et al., 2009; Scammacca et al., 2007; Slavin and Madden, 2011).

The idea that pedagogical factors are malleable is well described in epistemological theories on education and teaching. Hopmann (2015) argues that pedagogical practices depend on "well-established, basic social patterns of the understanding of schooling that have sedimented in the respective traditions" (p.18). In the formation of these patterns, students themselves are predicted to play an important role (Carlgren and Klette, 2008; Liljenberg, 2016; Pietarinen et al., 2017; Priestley et al., 2015) . International research uses the concept of "experienced curriculum" to capture the shapings of students' conceptions or beliefs (Olteanu and Olteanu, 2013). However, no systematic approaches have been developed to take the malleability into account in effect studies. In Qvortrup and Lykkegaard (2024a), we suggest an approach to investigate malleable factors based on a constructivist stance and two key points of constructivism: *subject-* and *context-dependency*. These dependencies refer to the understanding that teaching efforts cannot be understood from an objective perspective. Linked to variations in preferences and to previous experiences and from these derived expectations, different individuals perceive teaching differently. We often talk about social desirability bias, i.e., respondents' answers reflect what they believe is morally or socially acceptable as a source of error, but we argue that social desirability influences are part of every classroom practice and that they are therefore crucial to consider in the construction and use of instruments, not as a source of error, but as a way to ensure precision and thus reliability. This is supported by research indicating that student surveys can be as reliable as and actually may provide more information than both achievement scores and observation protocols, because students' *experiences* of teachers' behaviour in different situations are more important than the behaviour in itself (Fauth et al., 2014; Kane and Staiger, 2012; Peterson et al., 2000;

Worrell and Kuterbach, 2001). Furthermore, the individuals' perceptions depend on the shared understandings of going to and doing school that develop through communicative negotiations in different contexts in a complex interplay between the physical and social environment, available resources, traditions, etc. (Qvortrup and Lykkegaard, 2024a). When we talk about shared understandings of going to school, we refer to such things as rituals and routines associated with being a student, that is ways of behaving in class, sitting at their desks, listening to the teacher, raising hands), while when we talk about doing school, we refer to certain actions and division of responsibilities between teacher and students associated with certain forms of classroom organizations such as blackboard teaching, group work, assignments and tests or exams or with various teaching activities such as for instance inquiry-based learning or project-based learning. We suggest that these negotiations result in shared understandings that to some extent are stable both situationally (in a classroom between a particular teacher and a particular group of students) and more generalized (across classrooms and educational contexts within a particular educational tradition). Thus, in our model (Figure 1) we distinguish between notions (situational constructs) and conceptions (more generalized constructs).

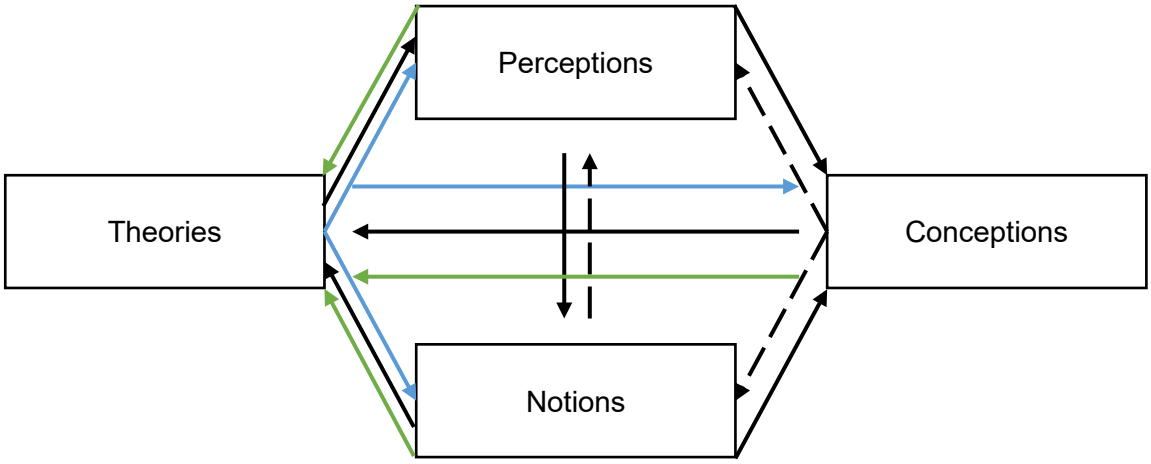

*Figure 1: The quartet of four constructs: theories and concepts (of learning, teaching etc.), contextual notions and subjective perceptions (*Qvortrup and Lykkegaard, 2024a*)*

In accordance with Qvortrup and Lykkegaard (2024a) we understand theories as systematic, worded and scientifically accepted descriptions (schemes) of educational matters such as connections between delimited purposes, objectives, and efforts. They are frames of expectations that guides teachers when "doing school." Conceptions are understood as generally applicable understandings of "going to school" (rules, rituals, and roles) and certain approaches to, or forms of teaching that are recognized as "doing school" in a school system or a given tradition. Notions are understood as situation-specific ways of "going to school" and "doing school," whereas perceptions refer to subjective experiences and expectations of what takes place in practice, that is, for concrete actualizations of "going to school" and "doing school." In Qvortrup and Lykkegaard (2024a) we show that notions and conceptions can be validly and reliably used in regression analysis to determine the effect of malleable factors.

Method and data

Participants

Survey data was collected as part of the QUINT project from 2,501 7th-9th grade students in five Nordic countries (Denmark (N=579), Norway (N=568), Sweden (N=537), Iceland (N=581), Finland (N=236)) and across 121 different classrooms with reference to three different courses (Social Science (N=678), Language Arts (N=897), and Mathematics (N=926), Finland has no data from Social Science classrooms). The survey responses had an equal distribution of boys and girls (Female (N=1,194), Male (N=1,197), missing (N=110)). We are not aware of whether there have been any common selection criteria and do not know whether the schools were deliberately chosen to broadly represent a cohort of classrooms, or whether they were randomly sampled. Thus, we cannot call on the data to be representative of the different countries, but as the intention is not to conclude anything about differences and similarities across the Nordic countries, but to investigate whether the seven key teaching characteristics differ across different Nordic contexts, we do not consider this to be a problem.

Survey

For the QUINT study, the 7C items were translated into Norwegian and pilot-tested on Norwegian students. This led to two additional items being added (Blikstad-balas and Roe, 2020). Six items are negatively phrased, these items caused confusion for some students in the Norwegian pilot study (Blikstad-balas and Roe, 2020), but all items were included in the QUINT study. The items were then re-translated into the additional Nordic languages. The item wording in Appendix 1 is our re-re-translation from the Norwegian and Danish items into English. This was accomplishable as written Norwegian and Danish are particularly similar to each other in vocabulary. The surveys were distributed to students in class and students were instructed to answer concerning the teacher in this specific class. Responses were given on a five-point Likert scale (1: Never, 2: Seldom, 3: Sometimes, 4: Often, and 5: Always). As students responded by pen and paper, some students ticked between the responses (e.g., between 3 and 4), for these responses a mean value is noted. Although responses are on a Likert scale, we assume an interval level of responses in the following analysis.

Analysis

We validate the instrument by following the steps for proper validation of educational instruments suggested by Schönrock-Adema et al. (2009). They suggest that, although the traditional way of validating instruments quantitatively adopting the original scales and examining the differences in scores on these scales between groups of respondents is a good method to establish construct validity, factor analysis should precede this kind of validation. Also, other researchers suggest that exploratory factor analysis is one of the most useful methods for studying and validating the internal structure of instruments (Ghazali et al., 2021; Henson and Roberts, 2006; Hoque and Awang, 2016; Kieffer, 1999).

Data preparation

Data were transferred to SPSS and all statistical analyses were conducted in SPSS v.28, SPSS AMOS v.28 and Excel.

Data were screened for impermissible values (no values were outside the 1-5 range) and subsequently screened for missing data (2 students had no item responses and were deleted). We checked for respondent misconduct and deleted 'unengaged respondents' with a standard deviation on responses lower than 0.25 for the 38 items (15 respondents were deleted). This resulted in a total of 2,484 respondents.

For the individual items, missing values ranged between 10 and 128 (0.4-5.2%). Since 623 students had missing data on one or more items (corresponding to 25% of the respondents), we imputed values for the missing data using series mean imputation. Previous research show that you can remedy up to 20-30% of missing data with an imputation technique and still obtain good parameter estimates (Eekhout et al., 2014; Hair, 2009).

Exploratory factor structure assessment

The dataset's suitability for factor analysis was evaluated using Kaiser-Meyer-Olkin measure of sampling adequacy (KMO > 0.5) and Bartlett's test of sphericity. Exploratory factor analysis (EFA) was used to openly investigate the underlying structure of the 7C items. Principal Axis Factoring was used, since the data did not meet multivariate normality assumptions (common for Likert scale items) and a Promax (oblique) rotation was chosen, as the factors were expected to correlate.

The attentive criteria

In a first round of exploratory factor analysis, we followed the guidelines and steps for proper use of factor analysis when validating educational environment instruments described by Schönrock-Adema et al. (2009). They recommend employing a combination of four psychometric criteria:

1. The scree test: the point of inflexion displayed by the scree plot (suggests 3 factors).
2. The eigenvalues criterion: in order not to overestimate the number of factors retained, factors with eigenvalues > 1.5 were accepted (suggesting 3 factors).
3. The variance explained criterion: a component was retained if it minimally explained an approximate additional 5% of the variance (suggesting 3 factors).
4. The interpretability or meaningfulness criteria: The interpretability of the factors was investigated using Hatcher's interpretability criteria (Hatcher, 1994):
    a. The rotated factor pattern should be a simple structure, meaning that items should load high on only one factor and low on the other (suggesting 3 factors).
    b. All factors should contain at least three items with significant loadings. The final set of items making up the factors was determined by iterative deleting items (cf. grey shading in Appendix 2) that did not load above 0.4 and the only two items loading on the third factor (suggesting two factors)
    c. Items loading on the same factor should share the same conceptual meaning and items loading on different factors should appear to measure different constructs. To determine the internal factors meaning, the two authors independently interpreted the factors and discussed their understanding (and naming). Based on this, we reached agreement on a two-factor model with the factors 'support' and 'control'.

According to Schönrock-Adema et al. (2009), this approach should yield interpretable and practically useful factors with less risk of under- or over-factoring and theoretically sensible dimensions that are more in accordance with educational theories.

The lax criteria
In a second round of EFA, we loosened the eigenvalue criteria above and used Kaiser's (1960) criteria (eigenvalues ≥ 1), moreover we dismissed the part of criteria 4b demanding that all factors should contain at least three items with significant loadings.

Country specific assessment
In a third and final round of EFAs, we conducted country specific factor analyses with the lax validation criteria described above.

Model fit assessment

Confirmatory factor analysis (CFA) was performed to investigate the model fit of the theoretical models as well as for the empirical exploratory models (items within the same factors were allowed to co-variance). The feasibility of the models for the current data set was tested by assessing model fit data: Chi-square, CMIN, CFI, RMSEA and SRMR (Bentler, 1990, 1992; Hu and Bentler, 1998).

Reliability and validity assessment

We evaluated the validity of the models by assessing construct reliability and construct validity.

Construct reliability

The internal consistency of the factors was tested using Cronbach's alpha (standardized $\alpha > 0.6$ (Ursachi et al., 2015)) and composite reliability (CR > 0.6, (Haji-Othman and Yusuff, 2022)).

Construct validity

In order to assess how well the items selected for the factors in each model actually measured the constructs, we assessed convergent validity (testing whether the items in one factor actually measures the same construct) and discriminant validity (testing whether the factors differ from each other).

For convergent validity we tested the average variance explained (AVE), since most values were below the cut off criteria AVE > 0.5 (the factor explains more than half of the variance in the belonging items), we used the cut off criteria AVE > 0.4 if composite reliability was good (CR > 0.6), (Fornell and Larcker, 1981).

For discriminant validity we used the heterotrait-monotrait ratio (HTMT < 0.85, (Kline, 2011)) to evaluate the degree to which the items that should not be very highly correlated with each other were distinct.

Results

For the analysis strategy presented above, we will first present the results of the validation on the total data set - i.e., across the Nordic contexts - after which we present the results of the validation within each of the Nordic contexts.

Assessing models suitable for the Nordic context

On the left side of Figure 2 below we represent the two theoretical models; the 3 domains model and the 7Cs model. Both of these models show acceptable construct reliability with Cronbach's' alpha > 0.6, and acceptable construct validity with CR > 0.6 except for the factor 'challenge' in the 7Cs model (see Appendix 2). However, both models have questionable construct validity as they each have a factor with poor average variance explained (AVE < 0.4 for CR > 0.6) and a pair of factors with poor heterotrait-monotrait ratio (HTMT > 0.85), see Appendix 4. Additionally both theoretical models have poor model fit (see Table 1). In summary, this shows evidence that the surveyed items are *not* adequately captured by these two theoretical models. We thus turn towards the exploratory empirical models (see right side of Figure 2).

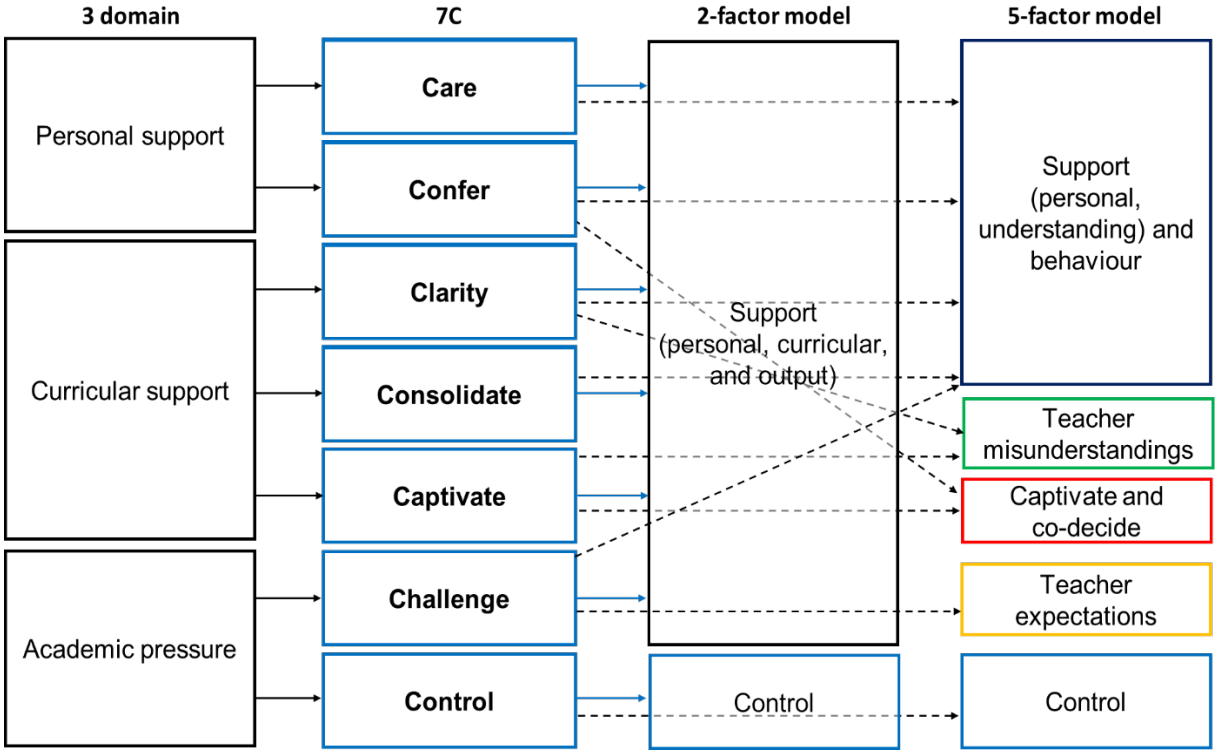

Figure 2: The four assessed models: The theoretical (3 domain and 7Cs) models on the left and the empirical models (the attentive 2-factor model and the lax 5 factor model) on the right. Arrows illustrate how items from 7C load on the empirical factors.

The exploratory 2 factor model (with the attentive criteria) has a better construct reliability (appendix 2), a better construct validity (appendix 4), and a better model fit than the two theoretical models (Table 1). The two factors are *Control* equalling the Control factor from the 7C framework and a very broad and multi-faceted factor consisting of all items from the 7C Care factor, two of five items from the Confer factor, four of six items from the Clarify factor, three of four items from the Captivate factor, all items from the Consolidate factor and two of seven items from the Challenge factor. The breadth and multi-facetedness of the factor make it hard to name, but our best suggestion is *Support (personal, curricular, and output)*. According to Schönrock-Adema et al. (2009), the attentive approach should yield interpretable and practically useful factors, and that criterion is by no means fulfilled. We therefore turn to the exploratory factor analysis with the softer set of criteria, which yielded five separate factors.

The five-factor model yields a moderate fit for the data (Table 1). The model has acceptable construct reliability (Appendix 2) and moderate construct validity (with one factor, the inverted questions with poor AVE, see Appendix 2), but with good HTMT (Appendix 4). The five factors are *Control* equaling the Control factor from the 7C and 4 factors that are all different from the factors in 7C. The first of these factors is a broad factor which is very much like, but not completely aligned with the 2-factor model's support factor. It consists of all items from the 7C Care and Consolidate factors, two of five items from the Confer factor about students' voices, three of six items from the Clarify factor about supporting student understanding, and one of seven items from the Control factor on student behavior. We term this factor *Support (personal, understanding) and behavior*. The second one is a factor consisting of one item from the 7C Confer factor and three of four items from the Captivate factor. We term this *Captivate and co-decide*, Finally, we have a factor consisting of three of six inverted questions about *teacher misunderstandings* and a factor consisting of two of seven items from the Challenge factor about teacher expectations. In sum, the five-factor model with the lax criteria was better than the theoretical models, but not as good as the 2-factor model (with the attentive criteria), but it had more interpretable and practically useful factors, although the *Support (personal, understanding) and behavior* factor is still quite broad. Reasoned in this and in the fact that it is discussable whether the Nordic countries are sufficiently similar to talk about the analysis being context-specific (Esping-Andersen, 1990; Telhaug et al., 2006), cf. our assumption about the context-specific malleability, we turn to the exploratory factor analysis with the softer set of criteria, on data divided according to national contexts.

Table 1: model fit data for the individual 9 models

	Chi-Square	**P**	**CMIN/df**	**CFI**	**RMSEA**	**Pclose**	**SRMR**
	(sample-size depen-dent)	Probabilit y level	The minimum discrepan -cy		The root mean square error of approxim a-tion	Test for the null hypothesi s for RMSEA	
Threshold values		**>0,05**	**<5,00**	**>0,90**	**<0,08**	**>0,05**	**<0,08**
3 domains	10.991,1	0,000	16,603	0,756	0,079	0,000	0,0868
7C	5.981,1	0,000	9,287	0,874	0,058	0,000	0,0584
2 Factors (attentive)	1160,9	0,000	3,922	0,974	0,034	1,000	0,040
5 Factors (lax)	2.299,1	0,000	5,540	0,945	0,043	1,000	0,0457
Denmark	592,4	0,000	1,834	0,955	0,039	1,000	0,0402
Norway	857,0	0,000	2,209	0,944	0,046	0,934	0,0501
Sweden	1148,6	0,000	2,558	0,918	0,054	0,038	0,0574

Finland	653,3	0,000	1,805	0,911	0,059	0,026	0,0595
Iceland	1005,6	0,000	2,585	0,927	0,053	0,129	0,0594

Assessing models suitable for individual Nordic countries

The country specific explorative models (Figure 3) have good construct validity (the models show acceptable Cronbach's' alpha (alpha > 0.6 except for one factor in the Danish model) and good construct validity (with CR > 0.6 for all items, except for two items in the Danish model), see Appendix 3. The construct validity is discussable (all models except for the Norwegian model have one factor (three for the Danish model) with poor average variance explained (AVE < 0.4 for CR > 0.6) but good heterotrait-monotrait ratio (HTMT < 0.85), see Appendix 4). Additionally, the models show good model fit (Table 1).

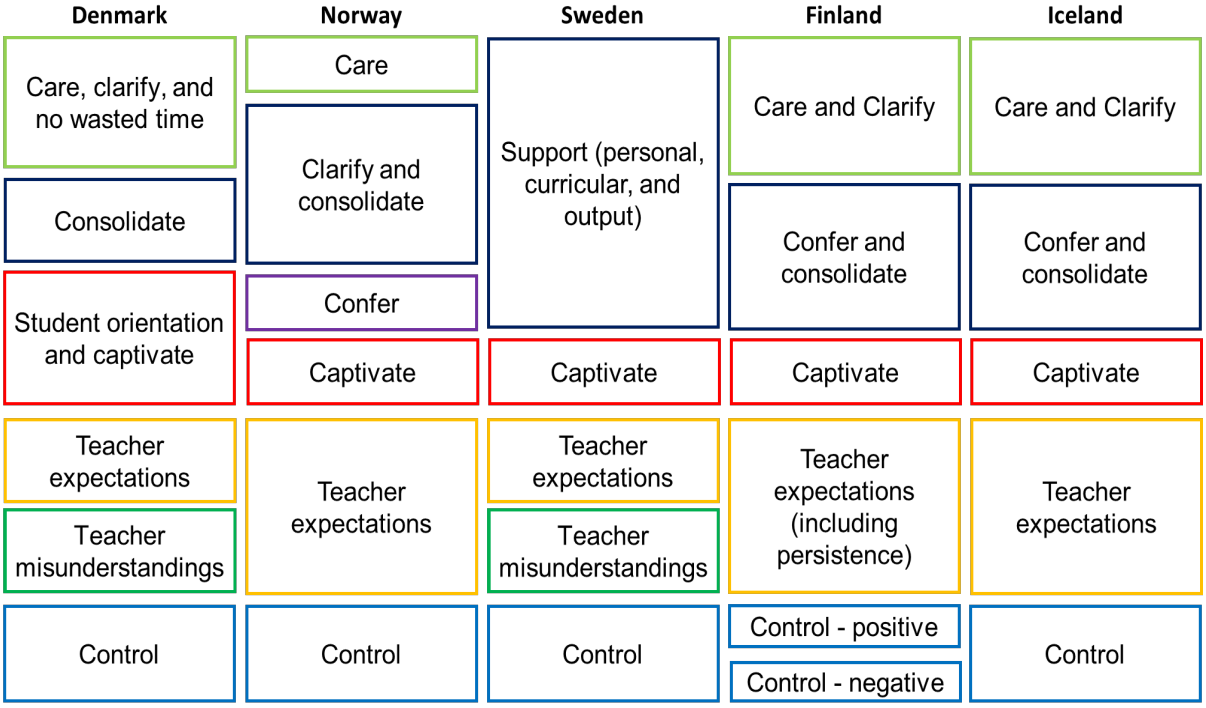

Figure 3: the five empirical country-specific models

As shown in figure 3, the models for Iceland and Sweden consist of five factors, while the Norwegian, Finnish, and Danish models consist of six factors. Although the items vary (cf. Appendix 3), we find what we can describe as a *Control* factor in all countries. However, it is noteworthy that it splits into two, respectively a negative and a positive control factor in Finland. In the four countries Iceland, Sweden, Norway, and Denmark, we find a completely similar *Teacher expectations* factor. In Finland the items of this factor are merged together with items on teacher's expectations regarding persistence, making it closer, but not aligned with, the 7C Challenge factor. In all countries, we find a factor consisting of the items from the 7C Captivate factor, except for one negatively worded item in Sweden, Norway, Finland, and Denmark. However, in Denmark these items are merged with three items from the Confer factor, and therefore the naming of this factor is slightly different here, where we choose *Student orientation and captivate* to cover the meaning of the factor. In Sweden, we find a broad support factor which is

very close to the one from the 2-factor model and, like there, is termed *Support (personal, curricular, and output)*. In the other countries, this factor is divided into *Care and clarify* and *Confer and consolidate* in Iceland and Finland (not completely the same across the two countries, but very similar), *Care* and *Clarify and consolidate* and *Confer* in Norway, and *Care, Clarify and no wasted time* and *consolidate* and the above mentioned *student orientation and captivate* in Denmark. Finally, it is interesting that – as we saw in the five-factor model above – negatively worded questions come together in one factor in two countries (Q13 and Q36 in Denmark, and Q13, Q36 and Q23 in Sweden), which we term *Teacher misunderstanding*.

Summarizing the results with reference to the theoretical figure (Figure 1), we can conclude that the theoretical expectations linked to the 7C framework were not adequately met in our empirical data, and that our analyses confirm the assumption of context specificity. It is beyond the scope of the article to examine the situation-specific context specificity (i.e. to examine notions or how the factors shape themselves at classroom level), and we have limited ourselves to focusing on nation-contexts, i.e. on conceptions (cf. Figure 1), but from the results we assume that the factors will look different at notion levels and that models at notion level will be even stronger.

Discussion

In the article we have considered nine factor structures: two theoretical models (the three domains model and the seven pedagogical factors model) and seven empirical models (the exploratory two-factor model, five-factor model, and five country-specific models with either five or six factors). None of the empirical models resulted in the 7C-factor structure anticipated from the 7Cs Framework. As described in the beginning of the article, other studies suggest that a 2-factor model and a 5-factor model predict the most variation in scores (Kuhfeld, 2016, 2017; Wallace et al., 2016). Our 2- and 5-factor models were similar but not completely aligned with these. Model fit, reliability, and validity assessment presented were better for the empirical models than the theoretical models. However, considerations of aspects besides overall fit indices e.g., adequacy and interpretability of parameter estimates, and model complexity remain critical in deciding on the validity of models.

Remarkably, the Control factor is consistent both across the different analyses in this article and in previous studies. Control – which in the 7C framework focuses on behaviour, respect, and the absence of wasting time (Appendix 1) – is thus a teaching characteristic that is recognized across many different educational contexts. This factor has been among the best predictors of other measures of quality (Ferguson and Danielson, 2015; Kane and Staiger, 2012), and thus is a quality indicator that is not context-dependent but stable across contexts. Behaviour, respect, and time utilization are thus perceived, notified, and conceptualised as important beyond national borders, which cannot be said to the same extent about the other pedagogical factors. It is remarkable in our analysis that the models and factor structures if we ignore this one control factor, vary so much across contexts.

As described in the theory section, the idea that pedagogical factors are malleable is well described in epistemological theories on education and teaching. We capture the idea of malleability due to subject and context dependencies in Figure 1's distinction between perceptions, notions, and conceptions. As

systematically described in Qvortrup and Lykkegaard (2024b) there has been no consistency between theories on epistemology and method development in the field. Thus, Howe (1988, 1992), Onwuegbuzie and Teddlie (2003) and Johnson and Onwuegbuzie (2004) note that there is often confusion about how to understand the relationship between epistemological paradigms and methodologies. There is a tendency to treat epistemological paradigms as "unnecessary edifice" (Stutchbury, 2022, p.113), to treat method and epistemology as being synonymous (Bryman, 1984; Howe, 1992) or to include an explanation of epistemological paradigms but failing to translate this into a practical method for gathering empirical data (Scott, 2010; Stutchbury, 2022). In addition to contributing to the validation of the 7Cs framework, we hope that our article, through its theoretical and empirical contributions, will inspire and invite further dialogue about and development of the methodological aspects. Our focus on the malleability of the factors and our identification of context-dependent factors challenges the possibility of creating generalizable evidence but increases in return the sensitivity to concrete contexts and also their applicability to developing quality in these contexts. We must become better at including pedagogical factors in effect studies because we can intervene on these factors and thus use them as a basis for developing quality in schools.

References

ARCHER, J., KERR, K.A., and PIANTA, R.C., (2014). *Why measure effective teaching*. In: T.J. KANE, K.A. KERR and R.C. PIANTA, eds., *Designing teacher evaluation systems: New guidance from the Measures of Effective Teaching Project.* San Francisco, CA: Jossey Bass. pp.1–5.

BENTLER, P.M., (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, **107**(2), p.238.

BENTLER, P.M., (1992). On the fit of models to covariances and methodology to the Bulletin. *Psychological Bulletin*, **112**(3), p.400.

BLIKSTAD-BALAS, M., and ROE, A., (2020). Hvordan opplever elevene norsktimene? In: M. BLIKSTAD-BALAS and A. ROE, eds., *Hva foregår i norsktimene? Utfordringer og muligheter i norskfaget på ungdomstrinnet* (Vol. 1). Universitetsforlaget.

BORMAN, G.D., HEWES, G.M., OVERMAN, L.T., and BROWN, S., (2003). Comprehensive School Reform and Achievement: A Meta-Analysis. *Review of Educational Research*, **73**(2), pp.125–230. https://doi.org/10.3102/00346543073002125

BRADSHAW, R., (2017). *Improvement in Tripod student survey ratings of secondary school instruction over three years*. Doctoral dissertation. Boston University.

BRYMAN, A., (1984). The debate about quantitative and qualitative research: a question of method or epistemology? *British Journal of Sociology*, **35**(1), pp.75–92. https://doi.org/10.2307/590553

CARLGREN, I., and KLETTE, K., (2008). Reconstructions of Nordic Teachers: Reform policies and teachers' work during the 1990s. *Scandinavian Journal of Educational Research*, **52**(2), pp.117–133. https://doi.org/10.1080/00313830801915754

CHAPLIN, D., GILL, B., THOMPKINS, A., and MILLER, H., (2014). Professional Practice, Student Surveys, and Value-Added: Multiple Measures of Teacher Effectiveness in the Pittsburgh Public Schools. REL 2014-024. Washington, DC: Regional Educational Laboratory Mid-Atlantic.

COHEN, J., RUZEK, E., and SANDILOS, L., (2018). Does teaching quality cross subjects? Exploring consistency in elementary teacher practice across subjects. *AERA Open*, **4**(3). https://doi.org/10.1177/2332858418794492

CROW, T., (2011). The view from the seats. *The Learning Professional*, **32**(6), p.24.

DE BOER, H., DONKER, A.S., and VAN DER WERF, M.P., (2014). Effects of the attributes of educational interventions on students' academic performance: A meta-analysis. *Review of Educational Research*, **84**(4), pp.509–545. https://doi.org/10.3102/0034654314540006

DETTERMAN, D.K., (2016). Education and intelligence: Pity the poor teacher because student characteristics are more significant than teachers or schools. *The Spanish Journal of Psychology*, **19**(E93), pp.1–11. https://doi.org/10.1017/sjp.2016.88

EDMONDS, M.S., VAUGHN, S., WEXLER, J., REUTEBUCH, C., CABLE, A., TACKETT, K.K., and SCHNAKENBERG, J.W., (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research*, **79**(1), pp.262–300. https://doi.org/10.3102/0034654308325998

EEKHOUT, I., DE VET, H.C., TWISK, J.W., BRAND, J.P., DE BOER, M.R., and HEYMANS, M.W., (2014). Missing data in a multi-item instrument were best handled by multiple imputation at the item score level. *Journal of Clinical Epidemiology*, **67**(3), pp.335–342. https://doi.org/10.1016/j.jclinepi.2013.09.009

ESPING-ANDERSEN, G., (1990). *The three worlds of welfare capitalism*. Princeton, NJ: Princeton University Press.

FAUTH, B., DECRISTAN, J., RIESER, S., KLIEME, E., and BÜTTNER, G., (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction,* **29**(1), pp.1–9. https://doi.org/10.1016/j.learninstruc.2013.07.001

FERGUSON, R.F., (2012). Can student surveys measure teaching quality? *Phi Delta Kappan*, **94**(3), pp.24–28.

FERGUSON, R F., and DANIELSON, C., (2015). How framework for teaching and tripod 7Cs evidence distinguish key components of effective teaching. In: T.J. KANE, K.A. KERR and R.C. PIANTA, eds., *Designing teacher evaluation systems: New guidance from the Measures of Effective Teaching Project.* San Francisco, CA: Jossey Bass. pp.98–143.

FERGUSON, R.F., and HIRSCH, E., (2014). How working conditions predict teaching quality and student outcomes. In: T.J. KANE, K.A. KERR and R.C. PIANTA, eds., *Designing teacher evaluation systems: New guidance from the Measures of Effective Teaching Project.* San Francisco, CA: Jossey Bass. pp.332–380.

FORNELL, C., and LARCKER, D.F., (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, **18**(1), pp.39–50. https://doi.org/10.1177/002224378101800104

GHAZALI, N., NORDIN, M.S., and TUNKU AHMAD, T.B., (2021). Development and Validation of Student's MOOC-Efficacy Scale: Exploratory Factor Analysis. *Asian Journal of University Education*, **17**(4), pp.327–339.

HAIR J.F., BLACK, W.C.,BABIN, B.J., and ANDERSON, R.E., (2009). *Multivariate Data Analysis*. Upper Saddle River, NJ: Prentice Hall Pearson.

HAJI-OTHMAN, Y. and YUSUFF, M.S.S., (2022). Assessing Reliability and Validity of Attitude Construct Using Partial Least Squares Structural Equation Modeling, *International Journal of Academic Research in Business and Social Sciences*, **12**(5), pp.378–385. https://doi.org/10.6007/IJARBSS/v12-i5/13289

HANUSHEK, E.A., (2011). The economic value of higher teacher quality. *Economics of Education review*, **30**(3), pp.466–479. https://doi.org/10.1016/j.econedurev.2010.12.006

O'ROURKE, N., and HATCHER, L., (2013). *A Step-By-Step Approach to Using SAS System for Factor Analysis and Structural Equation Modeling*. Cary, NC: SAS Institute.

HENSON, R.K., and ROBERTS, J.K., (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, **66**(3), pp.393–416. https://doi.org/10.1177/0013164405282485

HIVER, P., SOLARTE, A.C.S., WHITESIDE, Z., KIM, C.J., and WHITEHEAD, G.E., (2021). The role of language teacher metacognition and executive function in exemplary classroom practice. *The Modern Language Journal,* **105**(2), pp.484–506. https://doi.org/10.1111/modl.12707

HOPMANN, S., (2015). 'Didaktik meets Curriculum' revisited: historical encounters, systematic experience, empirical limits. *Nordic Journal of Studies in Educational Policy*, **2015**(1). https://doi.org/10.3402/nstep.v1.27007

HOQUE, A., and AWANG, Z., (2016). *Exploratory factor analysis of entrepreneurial marketing: Scale development and validation in the SME context of Bangladesh*. Proceedings of the International Social Sciences and Tourism Research Conference.

HOWE, K R., (1988). Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher*, **17**(8), pp.10–16. https://doi.org/10.3102/0013189X017008010

HOWE, K.R., (1992). Getting over the quantitative-qualitative debate. *American Journal of Education*, **100**(2), pp.236–256. https://psycnet.apa.org/doi/10.1086/444015

HU, L-T., and BENTLER, P.M., (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, **3**(4), pp.424–453. https://psycnet.apa.org/doi/10.1037/1082-989X.3.4.424

JOHNSON, R.B., and ONWUEGBUZIE, A.J., (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, **33**(7), pp.14–26. https://www.jstor.org/stable/3700093

JOHNSON, S.M., (2006). *The Workplace Matters: Teacher Quality, Retention, and Effectiveness*. Working Paper. National Education Association Research Department.

KAISER, H.F., (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, **20**(1), pp.141–151. https://doi.org/10.1177/001316446002000116

KANE, T.J., and STAIGER, D.O., (2012). *Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains*. Research Paper. MET Project. Bill & Melinda Gates Foundation.

KIEFFER, K.M., (1999). An Introductory Primer on the Appropriate Use of Exploratory and Confirmatory Factor Analysis. *Research in the Schools*, **6**(2), pp.75–92.

KLINE, R.B., (2011). Principles and Practice of Structural Equation Modeling. New York: Guilford Press.

KUHFELD, M.R., (2016). *Multilevel item factor analysis and student perceptions of teacher effectiveness*. Doctoral dissertation. University of California.

KUHFELD, M.R., (2017). When students grade their teachers: A validity analysis of the Tripod student survey. *Educational Assessment*, **22**(4), pp.253–274. https://doi.org/10.1080/10627197.2017.1381555

LILJENBERG, M., (2016). Teacher leadership modes and practices in a Swedish context–a case study. *School leadership & management*, **36**(1), pp.21–40. https://doi.org/10.1080/13632434.2016.1160209

OLTEANU, C., and OLTEANU, L., (2013). Enhancing mathematics communication using critical aspects and dimensions of variation. *International Journal of Mathematical Education in Science and Technology*, **44**(4), pp.513–522. https://doi.org/10.1080/0020739X.2012.742153

ONWUEGBUZIE, A.J., and TEDDLIE, C., (2003). A framework for analyzing data in mixed methods research. In: A. TASHAKKORI and C. TEDDLIE, eds., *Handbook of mixed methods in social and behavioral research.* Thousand Oaks, CA: Sage. pp.351–383.

PETERSON, K.D., WAHLQUIST, C., and BONE, K., (2000). Student surveys for school teacher evaluation. *Journal of Personnel Evaluation in Education*, **14**, pp.135–153. https://doi.org/10.1023/A:1008102519702

PHILLIPS, S.F., FERGUSON, R.F., and ROWLEY, J.F.S., (2021). Do they see what I see? Toward a better understanding of the 7Cs framework of teaching effectiveness. *Educational Assessment*, **26**(2), pp.69–87. https://doi.org/10.1080/10627197.2020.1858784

PIETARINEN, J., PYHÄLTÖ, K., and SOINI, T., (2017). Large-scale curriculum reform in Finland–exploring the interrelation between implementation strategy, the function of the reform, and curriculum

coherence. *The Curriculum Journal*, **28**(1), pp.22–40. https://doi.org/10.1080/09585176.2016.1179205

POLIKOFF, M.S., (2015). The stability of observational and student survey measures of teaching effectiveness. *American Journal of Education*, **121**(2), pp.183–212. https://doi.org/10.1086/679390

PRIESTLEY, M., BIESTA, G., PHILIPPOU, S., and ROBINSON, S., (2015). The teacher and the curriculum: Exploring teacher agency. In: D. WYSE, L. HAYWARD, and J. PANDYA, eds., *The SAGE handbook of curriculum, pedagogy and assessment*. New York: Sage. pp.187–201.

QVORTRUP, A., and LYKKEGAARD, E., (2024a). Malleable factors in teaching: why and how to address them from a constructivist perspective. *Scandinavian Journal of Educational Research*, **68**(3), pp.355–370. https://doi.org/10.1080/00313831.2022.2148272

QVORTRUP, A., and LYKKEGAARD, E., (2024b). Justifying methodologies in educational research: Epistemological paradigms as logics of justification and frames of inquiry. *Education Sciences*. Under review.

RIVKIN, S.G., HANUSHEK, E.A., and KAIN, J.F., (2005). Teachers, schools, and academic achievement. *Econometrica*, **73**(2), pp.417–458. https://doi.org/doi:10.1111/j.1468-0262.2005.00584.x

ROCKOFF, J.E., JACOB, B.A., KANE, T.J., and STAIGER, D.O., (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy*, **6**(1), pp.43–74. https://doi.org/10.1162/EDFP_a_00022

ROWLEY, J.F.S., PHILLIPS, S.F., and FERGUSON, R.F., (2019). The stability of student ratings of teacher instructional practice: examining the one-year stability of the 7Cs composite. *School Effectiveness and School Improvement,* **30**(4), pp.549–562. https://doi.org/10.1080/09243453.2019.1620293

SCAMMACCA, N., ROBERTS, G., VAUGHN, S., EDMONDS, M., WEXLER, J., REUTEBUCH, C.K., and TORGESEN, J.K., (2007). *Interventions for adolescent struggling readers: A meta-analysis with implications for practice*. Portsmouth, NH: RMC Research Corporation, Centre on Instruction. Available: http://files.eric.ed.gov/fulltext/ED521837.pdf

SCHEERENS, J., (2014). *Effectiveness of time investments in education: Insights from a review and meta-analysis.* New York: Springer. https://doi.org/10.1007/978-3-319-00924-7

SCHEERENS, J., (2017). The perspective of “limited malleability” in educational effectiveness: treatment effects in schooling. *Educational research and evaluation*, **23**(5-6), pp.247–266. https://doi.org/10.1080/13803611.2017.1455286

SCHEERENS, J., and MARKS, G.N., (2017). Malleability in educational effectiveness: what are realistic expectations about effect sizes? *Educational research and evaluation*, **23**(5-6), pp.143–147. https://doi.org/10.1080/13803611.2017.1455280

SCHWEIG, J., (2014). Cross-level measurement invariance in school and classroom environment surveys: Implications for policy and practice. *Educational Evaluation and Policy Analysis*, **36**(3), pp.259–280. https://doi.org/10.3102/0162373713509880

SCHÖNROCK-ADEMA, J., HEIJNE-PENNINGA, M., VAN HELL, E.A., and COHEN-SCHOTANUS, J., (2009). Necessary steps in factor analysis: enhancing validation studies of educational instruments. The PHEEM applied to clerks as an example. *Medical Teacher*, **31**(6), pp.226–232. https://doi.org/10.1080/01421590802516756

SCOTT, D., (2010). Education, epistemology and critical realism. London: Routledge.

SLAVIN, R., and MADDEN, N.A., (2011). Measures inherent to treatments in program effectiveness reviews. *Journal of Research on Educational Effectiveness*, **4**(4), pp.370–380. https://doi.org/10.1080/19345747.2011.558986

STUTCHBURY, K., (2022). Critical realism: an explanatory framework for small-scale qualitative studies or an 'unhelpful edifice'? *International Journal of Research & Method in Education*, **45**(2), pp.113–128. https://doi.org/10.1080/1743727X.2021.1966623

TELHAUG, A.O., MEDIÅS, O.A., and AASEN, P., (2006). The Nordic Model in Education: Education as part of the political system in the last 50 years. *Scandinavian Journal of Educational Research*, **50**(3), pp.245–283. https://doi.org/10.1080/00313830600743274

TIMMERMANS, A.C., (2012). Value added in educational accountability: Possible, fair and useful? Doctoral Dissertation. University of Groningen.

URSACHI, G., HORODNIC, I.A., and ZAIT, A., (2015). How reliable are measurement scales? External factors with indirect influence on reliability estimators. *Procedia Economics and Finance*, **20**, pp.679–686. https://doi.org/10.1016/S2212-5671(15)00123-9

WALLACE, T.L., KELCEY, B., and RUZEK, E., (2016). What can student perception surveys tell us about teaching? Empirically testing the underlying structure of the tripod student perception survey. *American Educational Research Journal*, **53**(6), pp.1834–1868. https://doi.org/10.3102/0002831216671864

WORRELL, F.C., and KUTERBACH, L.D., (2001). The use of student ratings of teacher behaviors with academically talented high school students. *Journal of Secondary Gifted Education*, **12**(4), pp.236–247. https://doi.org/10.4219/jsge-2001-670

YIK, B.J., RAKER, J.R., APKARIAN, N., STAINS, M., HENDERSON, C., DANCY, M.H., and JOHNSON, E., (2022). Evaluating the impact of malleable factors on percent time lecturing in gateway chemistry, mathematics, and physics courses. *International Journal of STEM Education*, **9**(1), pp.15–38. https://doi.org/10.1186/s40594-022-00333-3

Appendix 1: The 7C framework and questionnaire items

Domains	Pedagogical factors	Items	Item no.
Personal support	*Care*	The teacher makes me feel that s/he really cares about/is interested in me	Q1
		If something is bothering me/I have a hard time I think the teacher will know/notice	Q2
		The teacher tries/makes an effort to understand how students are doing	Q3
		The teacher is understanding when the students are tired or have had a long day	Q37
		The teacher takes time to help all students	Q38
	Confer	The teacher wants us to share our thoughts (and ideas) with each other	Q27
		The students have a say about how the various activities are done in this class	Q28
		The teacher gives us time to explain thoughts and ideas we have	Q29
		The students like to speak up about what they think of the class work	Q30
		The teacher respects my ideas and suggestions	Q31
Curricular support	*Captivate*	I don't have the strength to keep up in class because I get bored	Q23
		The teacher makes it fun to learn	Q24
		The teacher makes classes interesting	Q25
		I like the way we learn in this class	Q26
	Clarify	If you don't understand something the teacher explains it in another way	Q11
		The teacher knows if the students understand	Q12
		When the teacher is teaching, he/she thinks we understand even though we don't	Q13
		The teacher has several good ways to explain the things we are supposed to learn	Q14
		The teacher explains difficult things clearly	Q15
		The teacher goes through the material too quickly	Q36
	Consolidate	The teacher takes the time to summarize what we learn each day	Q32
		The teacher checks whether we understand what he/she has taught us	Q33
		We get useful comments so that we understand what we did wrong on assignments	Q34
		The comments that I get on my work in the class help me understand how to improve	Q35
Academic pressure	*Challenge*	When the students answer questions from the teacher, he/she asks them to explain or elaborate their answers	Q16
		The teacher expects our full effort/that we do our best in class	Q17
		The teacher doesn't let anyone give up, although the tasks are hard	Q18
		The teacher goes through the material too quickly	Q19
		The teacher wants me to justify my answers and explain why I think what I think	Q20
		In class, we learn a lot almost every day	Q21
		In this class, we learn to correct our mistakes	Q22
	Control	The students are behaving well in this class	Q4
		I don't like the way that students behave in this class	Q5
		The behaviour of the class annoys the teacher	Q6
		The behaviour of the class is a problem	Q7
		The class behaves the way the teacher wants them to	Q8
		The class treats the teacher with respect	Q9
		In class we work well and don't waste time	Q10

Appendix 2: Reliability and average variance explained for the models in Fig. 2

Factor structure reliability (Cronbach's alpha and construct reliability (CR)) and average variance explained (AVE) for the theoretical models: 3 domains and 7Cs and the explorative models with attentive and lax criteria respectively.

	3 domains			**7Cs**							**Total (attentive criteria)**		**Total (lax criteria)**				
Items	Personal support	Curricular support	Academic pressure	Care	Confer	Clarify	Captivate	Consolidate	Challenge	Control	Support (personal, curricular, and output)	Control	Support (personal, understanding) and	Control	Captivate and co-decide	Teacher misunderstandings	Teacher expectations
Q1	x			x							x		x				
Q2	x			x							x		x				
Q3	x			x							x		x				
Q37	x			x							x		x				
Q38	x			x							x		x				
Q27	x				x												
Q28	x				x						x				x		
Q29	x				x						x		x				
Q30	x				x								x				
Q31	x				x												
Q11		x				x					x		x				
Q12		x				x					x		x				
Q13 (inverted)		x				x										x	
Q14		x				x					x		x				
Q15		x				x					x						
Q36 (inverted)		x				x										x	
Q23 (inverted)		x					x									x	
Q24		x					x				x				x		
Q25		x					x				x				x		
Q26		x					x				x				x		
Q32		x						x			x		x				
Q33		x						x			x		x				
Q34		x						x			x		x				
Q35		x						x			x		x				
Q16			x						x								
Q17			x						x								x
Q18			x						x								
Q19			x						x								
Q20			x						x								x
Q21			x						x		x						
Q22			x						x		x		x				

Q4			x							x		x		x				
Q5 (inverted)			x							x		x		x				
Q6 (inverted)			x							x		x		x				
Q7 (inverted)			x							x		x		x				
Q8			x							x		x		x				
Q9			x							x		x		x				
Q10			x							x		x		x				
AVE (>0,5)	0,401	0,415	0,261	0,500	0,398	0,442	0,609	0,488	0,239	0,435	0,425	0,402	0,412	0,411	0,620	0,378	0,570	
CR (>0.6)	0,867	0,905	0,812	0,832	0,763	0,817	0,856	0,791	0,630	0,840	0,936	0,817	0,912	0,824	0,863	0,646	0,726	
Cronbach's alpha (>0,6)	0,863	0,902	0,783	0,822	0,759	0,809	0,841	0,786	0,653	0,836	0,938	0,836	0,912	0,836	0,852	0,646	0,726	

Appendix 3: Reliability and average variance explained for the models in Fig. 3

Factor structure reliability (Cronbach's alpha and construct reliability (CR)) and average variance explained (AVE) for the explorative models for the Danish, Norwegian, Swedish, Finnish, and Icelandic sample respectively.

	Denmark						Norway						Sweden					Finland						Iceland				
Items	Care, clarify, and no wasted time	Student orientation and captivate	Consolidate	Control	Teacher expectations	Teacher misunderstanding	Clarify, consolidate	Control	Captivate	Care	Confer	Teacher expectations	Support (personal, curricular, and output)	Teacher misunderstanding	Control	Captivate	Teacher expectations	Care and clarify	Confer and consolidate	Control - positive	Captivate	Teacher expectations (including persistence)	Control - negative	Care and clarify	Control	Confer and consolidate +	Captivate	Teacher expectations
Q1	x									x			x					x						x				
Q2	x									x			x					x						x				
Q3	x									x			x					x						x				
Q37													x					x							x			
Q38							x						x					x						x				
Q27					x						x								x						x			
Q28		x									x								x						x			
Q29		x									x		x						x						x			
Q30											x								x									
Q31		x											x					x							x			
Q11	x						x						x					x						x				
Q12	x						x						x					x						x				
Q13 (inverted)						x	x							x														
Q14	x						x						x											x				
Q15	x						x						x					x						x				
Q36 (inverted)						x	x							x														
Q23 (inverted)														x													x	
Q24		x							x							x					x						x	
Q25		x							x							x					x						x	
Q26		x							x							x					x						x	
Q32			x										x						x						x			
Q33			x				x						x						x						x			
Q34			x				x						x												x			
Q35			x				x						x												x			
Q16													x															
Q17					x							x					x					x						x
Q18																						x						
Q19																						x						
Q20					x							x					x					x						x
Q21																											x	
Q22			x				x						x															
Q4				x				x							x					x				x				
Q5 (inverted)								x						x								x		x				
Q6 (inverted)				x				x						x								x		x				
Q7 (inverted)				x				x						x								x		x				
Q8				x				x							x					x				x				
Q9								x							x					x				x				
Q10	x							x							x					x								

AVE (>0,5)	0,425	0,445	0,374	0,455	0,312	0,408	0,462	0,451	0,790	0,641	0,407	0,562	0,481	0,342	0,485	0,715	0,501	0,523	0,426	0,531	0,758	0,371	0,422	0,480	0,474	0,395	0,537	0,583
CR (>0.6)	0,853	0,827	0,748	0,767	0,561	0,574	0,916	0,851	0,919	0,843	0,733	0,719	0,940	0,608	0,789	0,882	0,661	0,907	0,814	0,819	0,903	0,694	0,681	0,880	0,841	0,851	0,849	0,736
Cronbach's alpha (>0,6)	0,864	0,856	0,756	0,788	0,629	0,562	0,897	0,847	0,916	0,843	0,729	0,718	0,939	0,802	0,784	0,875	0,643	0,906	0,813	0,814	0,670	0,692	0,670	0,895	0,854	0,852	0,839	0,734

Appendix 4: Discriminant validity

Model	Factors	HTMT
3 domains	Personal support - curricular support	0,92
	Personal support - academic pressure	0,72
	Curricular support - academic pressure	0,77
7C	Care - Confer	0,82
	Care - Clarify	0,87
	Care - Captivate	0,76
	Care - consolidate	0,84
	Care - Challenge	0,71
	Care - Control	0,46
	Confer - Clarify	0,75
	Confer - Captivate	0,74
	Confer - Consolidate	0,82
	Confer - Challenge	2,39
	Confer - Control	0,42
	Clarify - Captivate	0,78
	Clarify - Consolidate	0,82
	Clarify - Challenge	0,75
	Clarify - control	0,53
	Captivate - Consolidate	0,70
	Captivate - Challenge	0,63
	Captivate - Control	0,38
	Consolidate - Challenge	0,83
	Consolidate - Control	0,41
	Challenge - Control	0,42
2-factor model (exp)	Support (personal, curricular, and output) - Control	0,49
5-factor model (exp)	Teacher expectations – support (personal, understanding) and behaviour	0,48
	Teacher expectations - Control	0,30
	Teacher expectations - Teacher misunderstandings	0,17
	Teacher expectations - Captivate and co-decide	0,31
	Support (personal, understanding) and behaviour- control	0,53
	Support (personal, understanding) and behaviour- Teacher misunderstandings	0,65
	Support (personal, understanding) and behaviour- Captivate and co-decide	0,82
	Control – Teacher misunderstandings	0,46
	Control – Captivate and co-decide	0,42
	Teacher misunderstandings – Captivate and co-decide	0,62
Danish model (exp)	Teacher expectations – consolidate	0,54
	Teacher expectations – control	0,35
	Teacher expectations – Care, clarify, and no wasted time	0,55
	Teacher expectations - Teacher misunderstanding	0,48
	Teacher expectations – Student orientation and captivate	0,47

	Consolidate – control	0,37
	Consolidate - Care, clarify, and no wasted time	0,82
	Consolidate – Teacher misunderstanding	0,45
	Consolidate – Student orientation and captivate	0,73
	Control - Care, clarify, and no wasted time	0,51
	Control – Teacher misunderstanding	0,52
	Control – Student orientation and captivate	0,52
	Care – Teacher misunderstanding	0,59
	Care – Student orientation and captivate	0,78
	Teacher misunderstanding – Student orientation and captivate	0,60
Norwegian model (exp)	Teacher expectations – Clarify and consolidate	0,50
	Teacher expectations – control	0,32
	Teacher expectations – Care	0,37
	Teacher expectations – captivate	0,44
	Teacher expectations – confer	0,53
	Clarify and consolidate – control	0,47
	Clarify and consolidate – Care	0,74
	Clarify and consolidate – captivate	0,80
	Clarify and consolidate – confer	0,72
	Control – Care	0,35
	Control – captivate	0,38
	Control – Confer	0,37
	Care – Captivate	0,67
	Care – Confer	0,72
	Captivate – Confer	0,69
Swedish model (exp)	Teacher misunderstanding - Teacher expectations	0,27
	Teacher misunderstanding – control	0,03
	Teacher misunderstanding- Captivate	0,45
	Teacher misunderstanding - Support (personal, curricular, and output)	0,37
	Teacher expectations – control	0,46
	Teacher expectations - captivate	0,27
	Teacher expectations - Support (personal, curricular, and output)	0,47
	Control – Captivate	0,27
	Control – Support (personal, curricular, and output)	0,29
	Captivate – Support (personal, curricular, and output)	0,75
Finnish model (exp)	Teacher expectations (including persistence) - control positive	0,28
	Teacher expectations (including persistence) - control negative	0,07
	Teacher expectations (including persistence) – care and clarify	0,35
	Teacher expectations (including persistence) – captivate	0,23
	Teacher expectations (including persistence) – confer and consolidate	0,31
	Control positive - control negative	0,63
	Control positive – care and clarify	0,60
	Control positive – captivate	0,47
	Control positive – confer and consolidate	0,55

	Control negative – care and clarify	0,21
	Control negative – captivate	0,11
	Control negative – confer and consolidate	0,25
	Care and Clarify – Captivate	0,80
	Care and Clarify – Confer and consolidate	0,81
	Captivate – Confer and consolidate	0,73
Icelandic model	Teacher expectations - Captivate	0,29
	Teacher expectations - control	0,10
	Teacher expectations – Care and clarify	0,38
	Teacher expectations – confer and consolidate	0,50
	Captivate – control	0,49
	Captivate – care and clarify	0,49
	Captivate – confer and consolidate	0,77
	Control – care and clarify	0,49
	Control – confer and consolidate	0,41
	Care and clarify – Confer and consolidate	0,84