Education Journal, 2022, Vol. 50 No. 2, pp. 123–150

Development of an Instrument Measuring Student-Teacher Caring Relationships Using Nel Noddings's Conception of Care

Te-Hsin CHANG*

Holistic Education Center
Tzu Chi University of Science and Technology

Suzanne E. GRAHAM

Department of Education
University of New Hampshire

Caring relationships between students and teachers in schools are essential to social, emotional, and cognitive development, but quantitative measures of these relationships have been limited by inadequate or incomplete articulation of underlying theoretical frameworks. The purpose of this study was to develop a theoretically strong and psychometrically sound survey instrument to investigate student-teacher caring relationship in educational settings. This article describes the development of a new instrument, the Caring Relationship Survey — Student and Teacher Versions, based on Nel Noddings's conception of care. The survey items assess different aspects of Noddings's conceptual framework, with four sub-constructs of care: Engrossment, Motivational Displacement, Reciprocity, and Attribution of Best Motive Consonant with Reality. Item development included expert reviews and cognitive interviews with teachers and students. Survey responses were then obtained from 772 students and 629 teachers. The full sample was randomly divided in half for sequential exploratory and confirmatory factor analyses (EFA and CFA). EFA with the first subsample informed refinement of item composites, and then CFA with the second subsample compared four hypothesized measurement models. Consistent with Noddings's perspective, results confirm multiple dimensions of both students' and teachers' perspectives on care but with one more dimension of care than theorized by Noddings. The

* Corresponding author: Te-Hsin CHANG (tehsin.chang@gmail.com)

estimated Cronbach's alphas range between .68–.83 for the five sub-constructs. Further studies of the instrument using more geographically and culturally diverse samples will inform understanding of the instrument's external validity.

Keywords: survey development; ethic of care; student-teacher relationships, social-emotional learning

Decades of research in developmental psychology and education confirms the importance of relationships between teachers and students for social, emotional, and cognitive development (Phillippo et al., 2017; Williford & Wolcott, 2015). An essential component of these relationships is care. Many educational researchers indicate that care is a fundamental element of teaching; good teachers care (Goldstein & Lake, 2000; Kemp & Reupert, 2012; Vogt, 2002), and many teachers believe caring is an important aspect of professionalism (Goldstein & Lake, 2003; Lee & Ravizza, 2008; Robson & Bailey, 2009). Furthermore, when students perceive that their teachers care for them, there are positive effects on their motivation to learn (Lewis et al., 2012), engagement in classes (Cothran & Ennis, 2000), social and emotional development (Hamre & Pianta, 2001; Hughes et al., 2001; Wentzel, 1997), student's well-being (Graham et al., 2016), and ultimately on learning outcomes (Meyers, 2009; Pianta et al., 2012).

Although caring is one of the most important aspects of positive student-teacher relationships (Graham et al., 2016), the construct of care is often treated of secondary importance or implicitly assumed in empirical investigations (Velasquez et al., 2013). Those studies that do directly investigate caring relationships between teachers and students suffer from both theoretical and/or methodological limitations (Phillippo et al., 2017). In some studies, caring is not clearly operationalized or there are varying definitions of care (King & Chan, 2011; Lewis et al., 2012; Tosolt, 2008; Wentzel, 1997). In other studies, the connection between items measuring care and the operationalized constructs is not explicitly presented or it is incomplete (Huffman, 2005; Newton et al., 2007; Teven & McCroskey, 1997). Such limitations raise important concerns about the validity of instruments used to measure care in educational settings.

The purpose of the current study was to develop a new instrument to measure caring relationships between students and teachers in schools. To overcome limitations related to inadequate or incomplete construct definitions and lack of clearly articulated theoretical frameworks in the development of prior instruments measuring care, an important starting point for this work was to identify a strong theoretical foundation for instrument

development. Of course, from a theoretical perspective, interest in elucidating care in human relationships is not a new endeavor. Notable theorists such as developmental psychologist Carol Gilligan (1982) and philosophers Milton Mayeroff (1971) and Nel Noddings (1984, 2003) have written about the ethics of care from different disciplinary perspectives for decades. Although both Gilligan and Mayeroff identified various elements of care ethics, from a measurement perspective, these perspectives were not specified clearly enough or distinct enough from other related constructs to connect with specific items on a quantitative measure. Noddings, in contrast, provided a thorough conceptual analysis of care while also addressing care from an ethic of care perspective (Noddings, 1984, 2003), and her definition formed the foundation for the instrument developed in this study.

Most importantly, from a measurement perspective, Noddings's conceptualization clearly defines four specific features of care while simultaneously embracing the integrity and the complexity of care. She theorizes that the one-caring's consciousness has two necessary characteristics: *Engrossment* and *Motivational Displacement*. Engrossment is the ability to bracket oneself (putting aside one's own agenda) to pay attention to the needs and goals of the cared-for. With Motivational Displacement, the one-caring takes on the cared-for's projects or aims (expressed needs and goals) as one's own to assist them in accomplishing their goals. A third element of Noddings's theory is *Attribution of Best Motive Consonant with Reality*, in which the one-caring assumes the cared-for is well intentioned or has good motives unless there is evidence on the contrary. For example, teachers may believe that their students are well-intentioned even when the students break rules or do something wrong. Of course, teachers need a good understanding of their students and their students' individual situations to make genuine attributions of students' *best motives*, ones they know to also be *consonant with reality* (Noddings, 1984, 2003). A final and critically important element of Noddings's perspective is *Reciprocity*, which means that the cared-for acknowledges — even with something as small as a smile — that they receive the one-caring's actions as caring. Both care giver and care receiver must contribute to the interaction for the relationship to be called a caring one; caring is not something that can be done unilaterally.

Because Noddings's definition of care is clear and concise (Diller, 1988), it can be more easily operationalized than other theoretical perspectives on care. Although many empirical researchers studying caring relationships in schools do cite Noddings's works (e.g., Ng et al., 2013; Tosolt, 2008; Wentzel, 1997), they tend to use Noddings's conception of caring as evidence supporting their arguments about the importance of studying care but do not

specifically connect Noddings's definition of care with their survey items. In addition, few scholars examine both students' and teachers' perspectives when discussing caring relationships (Velasquez et al., 2013).

The Current Study

There is a clear gap in empirical research connecting specific, well-defined and clearly articulated theoretical frameworks to the development of survey instruments measuring caring relationships between students and teachers. This article directly addresses this limitation of prior research by describing the development and validation of a new instrument to measure caring relationships between students and teachers. The Caring Relationship Survey (CRS) — Student Version and Teacher Version were designed to better understand care in classrooms and thus help fill the gap in literature connecting care theory to empirical, quantitative measurement. Two versions of the instrument were developed: one examining students' perceptions and the other investigating teachers' perceptions of care in school. In contrast with prior measures of student-teacher caring relationships, this new survey has a strong theoretical foundation, explicitly linked with Noddings's (1984, 2003) important contemporary conception of care. This article describes the development of the specific items in the CRS, methods used to investigate its reliability and validity, and directions for future research.

Methods

Development of the Caring Relationship Survey (CRS)

The CRS — Student Version and Teacher Version were developed following general processes for scale development described by Carpenter (2018), DeVellis (2003), and Gehlbach and Brinkworth (2011). Survey development included the following: (a) identifying the theoretical concept(s); (b) item pool development and determining measurement format; (c) expert review of items; (d) cognitive interviews; (e) survey administration; (f) evaluating validity and reliability; and (g) scale modification.

As discussed above, Noddings's conception of care was chosen as the theoretical framework underlying survey development. Survey items were developed to measure the four major sub-constructs operationalized in Noddings's theory: (a) Engrossment; (b) Motivational Displacement; (c) Reciprocity; and (d) Attribution of Best Motive Consonant

with Reality. To directly acknowledge the relational aspect of caring, survey instruments were developed for both teachers and students.

Prior measures of care were used to develop an extensive initial item pool (Huffman, 2005; King & Chan, 2011; Tosolt, 2008). The strengths and weaknesses of each survey were taken into account in the current study. Huffman's (2005) survey directly referenced and employed Noddings's conception of care, but the items did not fully capture key elements of Noddings's theory. For example, Huffman's survey did not address relational aspects of caring relationships. In contrast, King and Chan's (2011) survey measured both students' and teachers' perceptions of care, but they did not employ a specific theory in their survey development. Finally, while Tosolt's (2008) survey items were informed by detailed review of literature on perceptions of care in schools, a specific theoretical framework of care was not articulated.

Taken together, these three instruments did provide a solid foundation to begin developing survey items to measure the complex construct of caring. Items consistent with Noddings's conception of care were identified and modified and then additional items were developed to fully encompass the different sub-constructs in Noddings's definition. Initial versions of both student and teacher versions of the CRS included 86 items. These items are statements of possible ways that teachers show caring behaviors in classrooms (e.g., *my teacher gives me second chances*), with respondents indicating their agreement on a 5-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = not sure, 4 = agree, 5 = strongly agree). The surveys include similarly worded items, with subject and verb differentiation related to the specific individuals (students or teachers) answering the questions.

The initial instruments were reviewed by three researchers with expertise in Noddings's conception of care. These experts examined connections between the items and the operationalized sub-constructs. Based on their feedback, the item pool was reduced from 86 to 50 items for both surveys.

Next, nine current and former K–12 teachers participated in cognitive interviews in which they were asked their interpretation of the teacher survey items. Items with ambiguous or confusing language were modified. These teachers also reviewed the student survey and provided suggestions about potential ways of clarifying the wording for the secondary school students who would be taking the survey. Finally, several secondary school students were asked to fill out the survey and provide feedback for additional item refinement.

At the end of this process, the student survey contained 42 items, and the teacher

survey contained 43 items. For the student survey, 12 items measure Engrossment, 17 items measure Motivational Displacement, 7 items measure Reciprocity, and 6 items measure Attribution of Best Motive Consonant with Reality. The teacher survey has an additional question on Reciprocity, with the same number of items as the student survey for the other sub-constructs. Two survey items were reverse-scored.

Participants and Procedures

Institutional Review Board's approval for Human Subjects was obtained prior to participant recruitment, data collection, and data analysis. The data was collected between December 2016 and August 2017. Participants included secondary school students and K–12 teachers and pre-service teachers from northern New England. Principals from six schools gave permission for us to administer the survey to students and teachers in their schools. In addition, email addresses of teachers were obtained from New England school websites and teachers were emailed invitations to participate in the study. Prior to administering surveys, parental consent, student assent, and teacher consent were obtained. The study was introduced to participants in their routine meetings or recruitment emails. When filling out the surveys, students and teachers were asked to keep a particular teacher and a class in mind respectively. Students completed paper-and-pencil surveys at their schools (n = 772). Teachers completed paper-and-pencil or online surveys (n = 629).

Demographic characteristics of participants are summarized in Table 1. The majority of the participants identified as White, consistent with state demographics for the schools in the sample. We acknowledge the lack of racial and ethnic diversity in this sample as a limitation of the research and discuss it further below.

Statistical Analyses

Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were sequentially conducted to examine the psychometric properties of both teacher and student versions of the CRS. Exploring the factor structure through EFA and then confirming the measurement model with CFA is a common process employed by survey researchers (Fabrigar & Wegener, 2012). Although the CRS was developed based on a strong theoretical framework, conducting EFA prior to CFA provided a more comprehensive approach to understanding the item loadings and thus a more nuanced understanding of different dimensions of students' and teachers' perceptions of care.

Table 1: Demographics of Teachers (*n* = 629) and Students (*n* = 772)

	Teachers									
	Grade level taught (*n* = 619)			Gender (*n* = 621)			Ethnicity (*n* = 618)		School type (*n* = 627)	
	Elem.	Middle	High	Male	Female	Others	White	Others	Private	Public
n	54	250	315	161	458	2	582	36	26	601
%	8.7	40.4	50.9	25.9	73.8	0.3	94.2	5.8	4.1	95.9

	Students								
	Grade level (*n* = 770)		Gender (*n* = 760)			Parent ed. level (*n* = 734)		Ethnicity (*n* = 763)	
	Middle	High	Male	Female	Others	Below College	College & above	White	Others
n	457	313	361	390	9	44	690	649	114
%	59.4	40.6	47.5	51.3	1.2	6.0	94.0	85.1	14.9

The full sample was randomly divided in half, resulting in one subsample of students (*n* = 386) and teachers (*n* = 314) for EFA, and a second subsample of students (*n* = 386) and teachers (*n* = 315) for CFA. Although a general concern of dividing any sample into smaller groups is potential effects on parameter estimate accuracy, these subsample sizes are sufficient for both EFA and CFA. For EFA, optimal participant-item ratios range from 3:1 to 20:1 (Williams et al., 2010). With 43 and 42 items on the teacher and student questionnaires respectively, sample sizes of 314 and 386 are acceptable. For CFA, Tomarken and Waller (2005) recommend a minimum sample size of 200 for four factor models. Descriptive statistics for the subsamples were compared to ensure that they were similar prior to conducting analyses.

Analyses of each of the datasets began with appropriate exploratory analyses. Univariate descriptive statistics were estimated for individual items and the composites measuring the four sub-constructs. Bivariate correlations were estimated between all items, with interpretation focusing on estimated correlations between pairs of items within each of the hypothesized composite. In addition, Cronbach's alphas were estimated to explore the internal consistency of the items in each group.

Exploratory factor analysis

To preliminarily understand the factor structure of the data, the first half of the dataset was analyzed using EFA. Missing data were handled by pair-wise deletion to maintain

the highest possible number of responses, and items were investigated to ensure that assumptions of EFA were met. EFA was initially conducted with the principal axis factoring method to determine the number of unique factors underlying the data. Alternative factoring methods were also considered. In addition, Kaiser-Meyer-Olkin measures of sampling adequacy were also estimated. Values closer to 1 suggest that items can be grouped into a smaller set of underlying factors. Finally, Bartlett's Test of Sphericity was conducted to determine potential relationships among the items. Factors with eigenvalues above 1.0 were retained and factor loadings reviewed. In addition, scree plots were reviewed to provide further information about the number of factors that should be retained (Fabrigar & Wegener, 2012). Cronbach's alphas of the resulting composites were estimated and item analysis was conducted to determine whether individual items could be removed from any composite to either increase or minimally decrease the estimated internal consistency.

Confirmatory factor analysis

Theory combined with the EFA results informed specification of four hypothesized CFA measurement models that were fitted to the second student and teacher datasets using Mplus (Muthén & Muthén, 2017). Model fit was evaluated using the χ^2 goodness-of-fit statistics for nested models. Schreiber et al. (2006) suggest that a ratio of χ^2 to *df* less than or equal to 2 or 3 represents a better fitting model. Model fit indices were also used to evaluate the goodness of fit: (a) Standardized Root Mean Squared Residual (SRMR $\leq$ .08); (b) Comparative Fit Index (CFI $\geq$ .95); (c) Tucker-Lewis Index (TLI $\geq$ .95); (4) Root Mean Squared Error of Approximation (RMSEA $\leq$ .06) (Hu & Bentler, 1999). The various measures of model fit were compared across the different fitted models to determine which best fit the empirical data.

Results

Descriptive Statistics

Table 2 contains descriptive statistics for the four originally hypothesized item composites from the subsamples of teachers and students subsequently used to conduct EFA and CFA. Descriptive statistics from both samples were similar, providing evidence that the two groups formed by random assignment were comparable. Average composite scores were higher for teachers than for students, suggesting that teachers perceive themselves as

Table 2： Descriptive Statistics for Composites

	EFA subsample					
	Teachers (*n* = 314)			Students (*n* = 386)		
Construct (items)	Range	Mean	*SD*	Range	Mean	*SD*
Engrossment (12)	3.75–5	4.44	0.30	1–5	3.69	0.61
Motivational Displacement (17)	3.59–5	4.30	0.33	1.24–5	3.63	0.60
Reciprocity (8/7)	2.88–5	4.09	0.44	1.29–5	3.37	0.63
Attribution of Best Motive (6)	2.33–5	3.97	0.48	1.67–5	3.69	0.60

	CFA subsample					
	Teachers (*n* = 315)			Students (*n* = 386)		
Construct (items)	Range	Mean	*SD*	Range	Mean	*SD*
Engrossment (9)	3.42–5	4.40	0.32	1.50–5	3.72	0.62
Motivational Displacement (8)	3.29–5	4.25	0.36	1.14–5	3.41	0.68
Reciprocity (8/7)	2.50–5	4.01	0.45	1.53–5	3.69	0.61
Attribution of Best Motive (5)	2.20–5	3.88	0.51	1.33–5	3.75	0.61

exhibiting higher levels of care than students perceive their teachers as caring. Additionally, smaller standard deviations for the teachers indicate more consistent responses compared with the students.

Estimated Bivariate Correlations and Internal Consistency

Prior to conducting EFA, correlations were estimated between items within composites for both teacher and student datasets. As expected, most items designed to measure the same underlying constructs are positively correlated. However, within the Motivational Displacement sub-construct for teachers, correlations are near to zero between one negatively worded item, #10, and two other items, #25 (r = –.04) and #35 (r = –.01). Further investigation of this item and a second negatively worded item (#33) is discussed below.

Cronbach's alphas were estimated to explore the internal consistency of the initial item composites (Table 3). With estimated alphas ranging between .71 to .89 before survey modification, the original item groups developed for the CRS have good reliability (DeVellis, 2003). However, even given good internal consistency, it is still useful to explore the possibility of alternative factor structures of the items with EFA (Fabrigar & Wegener, 2012).

Table 3: Estimated Cronbach's Alphas of Item Composites From Subsamples Used for EFA and CFA, Before and After Modifications

	Before modifications			After modifications				
		EFA subsamples			EFA subsamples		CFA subsamples	
		Teachers (n = 314)	Students (n = 386)		Teachers (n = 314)	Students (n = 386)	Teachers (n = 315)	Students (n = 386)
Composite	No. of items	Alpha	Alpha	No. of items	Alpha	Alpha	Alpha	Alpha
Engrossment	12	.78	.85	9	.73	.81	.75	.82
Motivational Displacement	17	.82	.89	8	.75	.81	.78	.83
Reciprocity	8/7	.75	.74	8/7	.78	.76	.77	.83
Attribution of Best Motive	6	.71	.74	5	.68	.71	.70	.73
Non-Traditional Roles	N/A	N/A	N/A	5	.67	.79	.68	.74

Exploratory Factor Analysis

EFA with principal axis factoring was conducted with the first subsamples of teachers (n = 314) and students (n = 386) to determine the number of unique aspects of caring underlying the survey data. Kaiser-Meyer-Olkin measures of sampling adequacy values are close to 1 (.895 for teachers and .940 for students) suggesting that the items can be grouped into a smaller set of underlying factors. Bartlett's Test of Sphericity (p < .001) for both groups indicates relationships among the items. These tests support conducting EFA for both teacher and student datasets.

The initial EFA resulted in 12 factors with eigenvalues above 1.0 for the teacher dataset and 8 for the student dataset. The first factor explains 24.9% of the variance for the teacher dataset and 32.6% of the variance for the student dataset. Review of factor loadings showed that 34 items from the teacher survey and 40 items from the student survey load on the first factor and have loadings greater than or equal to .40. This loading pattern is consistent with the theory that care is an overarching construct with possibly lower-level sub-constructs. To further explore this possibility, a second EFA with the Direct Oblimin rotation was employed, appropriate when one hypothesizes an overarching construct with correlated sub-constructs (Fabrigar & Wegener, 2012). The resulting Pattern Matrices revealed a more distinct factor loading pattern for both datasets. Although 12 eigenvalues were above 1.0 for the teacher data and 8 for the student data, a review of scree plots suggested four potential

factors for the teachers' responses and three for the students. These results are generally consistent with Noddings's theorized four sub-constructs, although there is one fewer factor than hypothesized by theory for the students. Additionally, since the EFA results suggest the possibility of a single factor for both datasets, this possibility was subsequently tested with CFA.

The results of the EFA also suggested the need for further exploration of the item groupings for both teacher and student datasets. Items not in their originally designed groupings and items with lower loading values were re-evaluated. If the only rationale for alternative item groupings was factor loadings over .40, the original theoretically informed item groups were instead retained. For example, the EFA of the teacher data did not group items #11 and #32 with other items designed to measure Attribution of Best Motive. However, since there was no theoretical rationale for including these items in different composites and they represent important aspects of Noddings's definition, both were retained in the Attribution of Best Motive composite.

Based on EFA results and theoretical rationale, eight items were removed from the teacher survey (see Table 4). Two of these items (#10 and #33) were reverse-coded items with negative loadings lower than .40. While including negatively worded items in a survey can help monitor whether participants are carefully reading the questions and avoid affirmation bias (DeVellis, 2003), it can also interrupt the flow of filling out the survey (Gehlbach & Brinkworth, 2011). The other eliminated items had low factor loadings and either: (a) had confusing wording (e.g., item #29); (b) were similar to other items (e.g., items #18, #37); or (c) did not load on any factors.

A final revision to the items on the teacher survey involved further investigation of two factors with five items that had negative loadings. It appeared that these items do not necessarily reflect traditionally recognized roles of teachers as cultivators of their students' academic learning and performance (e.g., items #34, #27). Instead, the items get at whether teachers believe that students' needs or interests *are more important* than their academic performance. These questions are clearly consistent with Noddings's notion of care because they challenge teachers to consider their perceptions of what it means to be a teacher. Noddings's conception of care building is a part of an ethic philosophy which encourages teachers to see students' needs and goals as important as teachers' tasks of teaching academic content. Therefore, we hypothesized that these items measure a unique dimension of care related to yet distinct from Noddings's other dimensions of care measured in the survey. We thus grouped these items together to measure a fifth sub-construct that we

labeled *Non-Traditional Roles* and which was subsequently tested in CFA models. Table 4 lists the original and modified items for each composite.

Table 4: Original and Revised Item Groups for the Teacher Survey; the student Survey Includes the Same Items (Except #43) With Wording Changed to Reflect the Student Perspective

Items originally designed to measure each construct		Items measuring each construct after modification					
		A	B	C	D	Fifth construct	Deleted items
A. Engrossment							
Q1.	I know how to keep my students on task.	✓					
Q2.	I know how to encourage my students.	✓					
Q4.	I pay attention to my students' needs in class.	✓					
Q5.	I pay attention to my students' feelings.	✓					
Q6.	I know my students' strengths and weaknesses.	✓					
Q12.	I pay attention to my students' academic achievements.	✓					
Q18.	I listen to my students' side of the story when conflict occurs.						×
Q20.	I take a personal interest in what my students do outside their classes.		✓				
Q21.	I listen to my students carefully and attentively when they are speaking.						×
Q22.	I pay attention to the spelling and pronunciation of my students' names.	✓					
Q23.	I know that sometimes my students do not want to do what the school expects them to do.	✓					
Q37.	I listen to stories from all sides before I decide what to do about a problem/ conflict.		✓				
B. Motivational Displacement							
Q3.	I give my students second chances.		✓				
Q9.	I see my students as individuals first, then as students.		✓				
Q10R.	I do not make time for my students when they need me.						×
Q13.	I know my students' background when they ask a question.	✓					
Q14.	I respond to my students' specific requests as much as I can.			✓			

Items originally designed to measure each construct		Items measuring each construct after modification					
		A	B	C	D	Fifth construct	Deleted items
Q16.	I am trying my best to respond to my students' requests.			✓			
Q24.	I work with my students to help them accomplish their academic goals.		✓				
Q25.	I work with my students to help them accomplish their goals outside the classroom.		✓				
Q26.	I make sure that my students understand the directions given in class.						×
Q27.	I consider my students' needs to be more important than my tasks as a teacher.					✓	
Q28.	I honor my students' opinions even when they are different from my own.						×
Q29.	I help my students to achieve their goals while also maintaining caring relationships with other students.						×
Q33R.	I do not treat my students the way they want to be treated.						×
Q35.	I work toward what is the best for my students even when it is not academically related.					✓	
Q36.	I give students opportunities to express their opinions on decisions that affect them.		✓				
Q38.	My students trust that I have their best interests in mind.					✓	
Q41.	I encourage my students to think they are intelligent.						×
C.	Reciprocity						
Q7.	I want to know if my students think that I am a caring teacher.			✓			
Q15.	I communicate with my students in order to know what they need.			✓			
Q17.	I look for my students' reactions to my care for them.			✓			
Q19.	I ask my students what I can do to help them reach their goals.		✓				
Q34.	I think what my students want to do is as important as what the school wants them to do.					✓	
Q40.	My students seek-out opportunity to talk			✓			

Items originally designed to measure each construct		Items measuring each construct after modification					
		A	B	C	D	Fifth construct	Deleted items
	to me.						
Q42.	My students do something to let me know that they receive my care.			✓			
Q43.	My students perceive me as a caring teacher. [This item is not on student survey]			✓			
D.	Best Motive Consonant with Reality						
Q8.	I trust that my students try to be good people.				✓		
Q11.	I believe my students can be good at what they want to do.				✓		
Q30.	I assume that my students have good intentions or motives for their behaviors.				✓		
Q31.	I believe that my students have good intentions even when they break the rules.				✓		
Q32.	I do not hold it against my students if they do something that they are not supposed to do.				✓		
Q39.	My students trust me to give them the benefit of the doubt.					✓	

Further investigation of the results from EFA of the student data showed somewhat different patterns of factor loadings, but the grouping of items informed by EFA was not as well connected with theory as it was for the teacher data. Therefore, we chose to proceed with CFA using consistent item groupings for teachers and students. The revised survey contains 35 items for the teacher survey and 34 items for the student survey: 9 Engrossment items; 8 Motivational Displacement items; 8/7 Reciprocity items; 5 Best Motive items; and 5 Non-Traditional Roles items.

Estimated Cronbach's alphas for the resulting item composites are presented in Table 3. Comparison of alpha reliability estimates before and after modifications to the item groups indicates that the estimated reliabilities of the new composites, while reasonably high, are slightly lower than the estimated alphas for the original groups. Such changes in internal consistency are to be expected when the number of items in an instrument are reduced (DeVellis, 2003), and the reduction in estimated reliability is outweighed by the advantage of developing a more parsimonious survey.

Confirmatory Factor Analysis

Grounded in Noddings's definition of care and informed by the EFA results, four CFA models were hypothesized and tested. These hypothesized models differ in two primary ways. First, Noddings theorizes that care is a multi-dimensional construct, and since it is possible that there is an overarching *care* construct encompassing the sub-constructs, we compared models containing only first-order factors with those including both first-order and second-order factors. Second, while Noddings's definition of care theorizes four primary sub-constructs, the results of the EFA discussed above suggested the possibility of a fifth dimension of care, Non-Traditional Roles. Therefore, the following four models were fitted to both the teacher and student data (see Appendix):

Model 1: In this model, a second-order factor of care was hypothesized, with five sub-constructs as first-order factors (Engrossment, Motivational Displacement, Reciprocity, Attribution of Best Motive, and Non-Traditional Roles).

Model 2: This model contains all five sub-constructs as first-order factors, with no second-order factor.

Model 3: In this model, as in Model 1, an overarching *care* construct is hypothesized as a second-order factor, but in contrast with Model 1, Model 3 contains four sub-constructs as first-order factors (Non-Traditional Roles removed).

Model 4: This model is like Model 3 (four sub-constructs of care as first-order factors), but does not include a second-order factor.

Table 5 presents the results of fitting the four hypothesized models to both the teacher and student datasets. The first conclusion to be drawn from reviewing the fit statistics from the models fitted to the teacher data is that Models 1 and 2 fit substantially better than Models 3 and 4, indicating that models including the new Non-Traditional Roles sub-construct are preferable. Comparison of Models 1 and 2 suggests a somewhat better fit for Model 2. While the difference in the χ^2 statistics of the two models favors Model 2 (difference in $\chi^2(5) = 18.104$, $p < .001$), other fit statistics are similar. The ratio of χ^2 to *df* is 2.53 for Model 1 and 2.52 for Model 2, indicating reasonable fit for both models. The CFI and TLI values for both models are closer to the recommended .95 than the CFI and TLI for Models 3 and 4. In addition, the RMSEA values for both models are slightly higher than the recommended .06, but the lower end of the 90% Confidence Interval is .065 for both fitted models. The SRMR is under .08 for both models. Since the goodness-of-fit statistics are slightly better for Model 2 compared with Model 1, and substantially better than those for

Table 5: CFA Model Fit Statistics

	Teacher data						
	χ^2 (df) χ^2/(df)	RMSEA [90% CI]	SRMR	CFI	TLI	AIC	BIC
Model 1	1401.811 (555) 2.53	.070 [.065, .074]	.073	.764	.747	19805.849	20087.292
Model 2	1383.707 (550) 2.52	.069 [.065, .074]	.072	.768	.749	19797.746	20097.952
Model 3	1886.542 (561) 3.36	.087 [.082, .091]	.146	.631	.608	20278.580	20537.508
Model 4	1883.551 (559) 3.37	.087 [.082, .091]	.146	.631	.607	20279.589	20546.022

	Student data						
	χ^2 (df) χ^2/(df)	RMSEA [90% CI]	SRMR	CFI	TLI	AIC	BIC
Model 1	1131.248 (522) 2.17	.055 [.051, .059]	.046	.889	.881	31413.529	31702.305
Model 3	1954.718 (528) 3.70	.084 [.080, .088]	.188	.741	.724	32224.999	32490.040
Model 4	1952.204 (526) 3.71	.084 [.080, .088]	.188	.741	.723	32226.485	32499.437

both Models 3 and 4, the best fitting CFA model for the teacher data includes five first-order factors, including Non-Traditional Roles, but no second-order factor.

Turning to the CFA results from the analyses of the student data, Model 2 did not converge so it is not included in the summary table. Consistent with the analysis of the teacher data, Model 1 is better fitting than either Model 3 (difference in $\chi^2(6) = 823.47$, $p < .001$) or Model 4 (difference in $\chi^2(4) = 820.956$, $p < .001$). The ratio of χ^2 to *df* of 2.17 also confirms that Model 1 fits better than Models 3 and 4. Other goodness-of-fit indices indicate reasonable fit for Model 1 with the RMSEA less than .06 and SRMR less than .08. However, the CFI and TLI, while both close to .90, do not achieve the recommended cutoffs value of .95. Taken together, the fit indices suggest a reasonable but not outstanding fit of Model 1 to the student data. As with the analysis of the teacher data, an important conclusion to be drawn from the analysis of the student data is the potential presence of a fifth sub-construct (Non-Traditional Roles).

The best fitting model for the student dataset is Model 1, whereas for the teacher dataset the best fitting model is Model 2. Figures 1 and 2 present standardized estimates for

Figure 1: Standardized Estimates From the Best Fitting Factor Model Underlying the Teacher Data (Model 2)

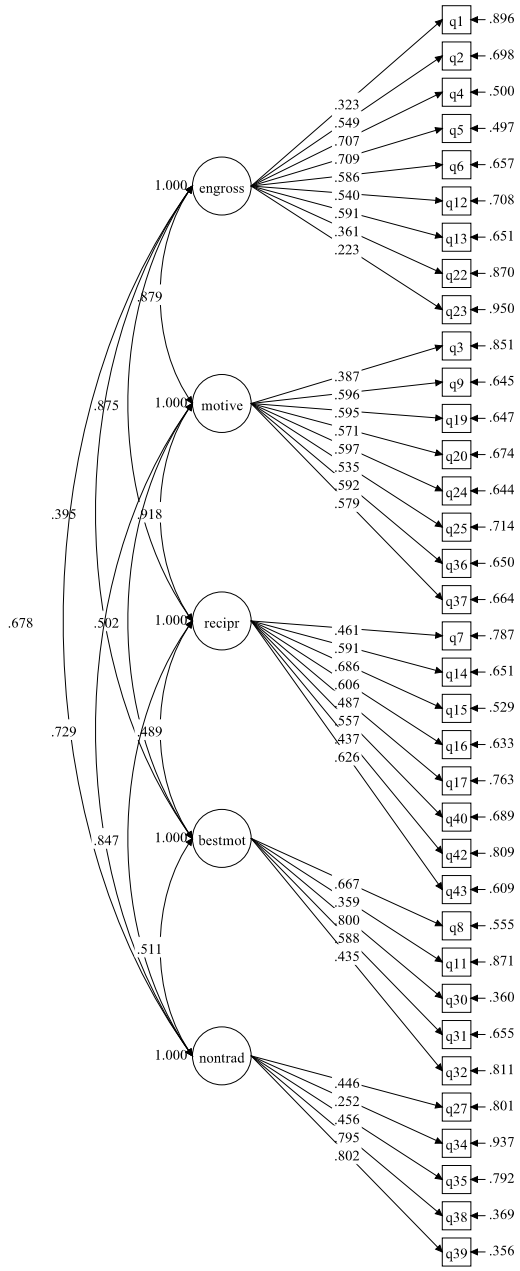
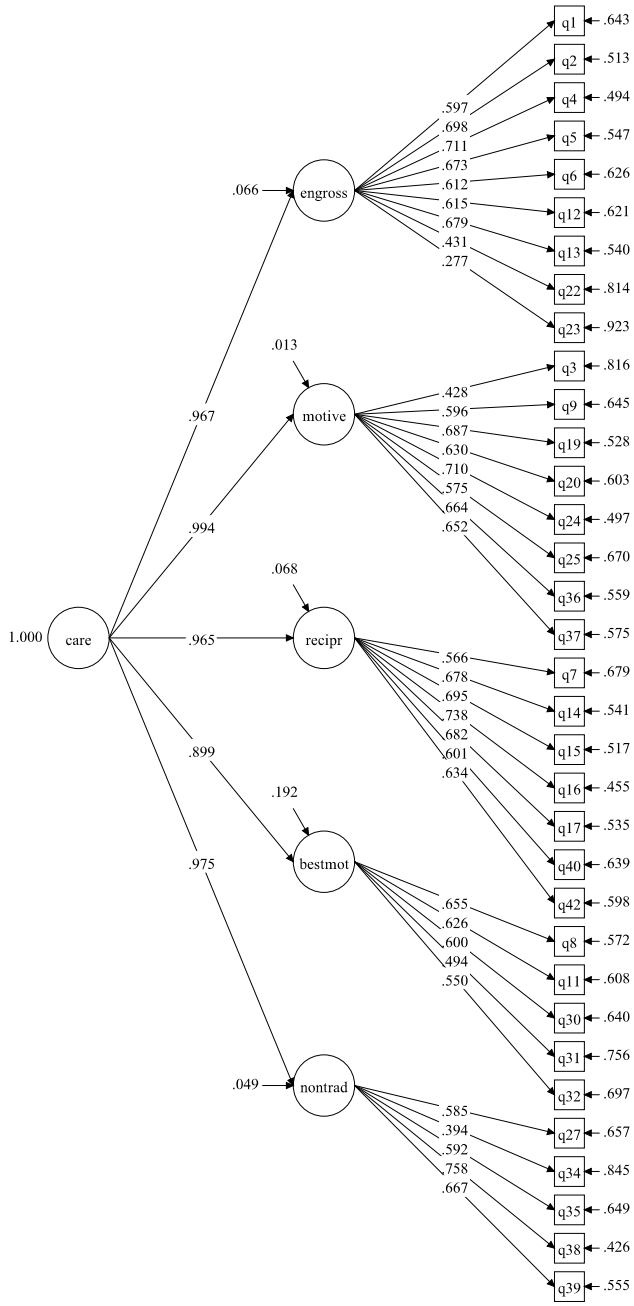


Figure 2: Standardized Estimates From the Best Fitting Factor Model Underlying the Student Data (Model 1)



these fitted models. For students, the five factors appear to capture a similar concept falling under the one umbrella term of care. For teachers, there is a clearer distinction among the five sub-constructs, although Model 1 was not a much worse fitting model than Model 2. We conclude that taken together the sub-constructs represent care as an overall concept.

Since some of the goodness-of-fit statistics of the best fitted models did not meet standard cut-off points, model modifications were considered. First, parameter estimates were investigated to potentially identify and remove non-significant paths, known as theory trimming (Kelloway, 2015). Although all paths are statistically significant, some of the parameter estimates are lower than .40. For example, #23 has small parameter estimates (.223 and .277 for analyses of the teacher and student data respectively), but both parameter estimates are statistically significant. Removing these items does not improve the overall goodness-of-fit indices for either the teacher or student models, so the items were retained.

The estimated correlations between sub-constructs are all very high for the student data (lower diagonal of Table 6). They are somewhat lower for the teacher data (upper diagonal of Table 6), particularly for estimated correlations involving Best Motive. In addition, the estimated correlations involving Non-Traditional Roles, while higher than .5, are lower for the teacher data than the student data. This is not surprising since the CFA of the teacher data suggests five distinct sub-constructs underlying care. The smaller correlations between the sub-constructs for the teachers suggest that although the sub-constructs are clearly related, they are also distinct (supporting our choice of Model 2 as the best model for the teacher data). For the student data, the high correlations between the sub-constructs support the existence of an overarching second-order construct of care for the students.

Cronbach's alphas were estimated for the items in the sub-constructs confirmed by the CFA analyses (last two columns of Table 3), and all are reasonably high. In addition, the estimated alpha values from the datasets used for CFA are slightly higher than from the datasets used for EFA. There is generally similar internal consistency for the item groups across the two samples, confirming that the composites have good reliability.

Table 6: Estimated Correlations Between Sub-constructs (Lower Diagonal for Student Data and Upper Diagonal for Teacher Data)

Observed variable	1	2	3	4	5
Engrossment	1	.879	.875	.395	.678
Motivational Displacement	.961	1	.918	.502	.729
Reciprocity	.939	.960	1	.489	.847
Attribution of Best Motive	.923	.859	.879	1	.511
Non-Traditional Roles	.929	.988	.923	.903	1

Discussion

The purpose of this study was to develop a new quantitative measure of caring relationships between students and teachers and to investigate its psychometric properties. We argue that researchers studying social-emotional learning often implicitly assume the importance of care without explicitly defining it and that studies of relationships between students and teachers have suffered from undertheorized concepts of care. The Caring Relationship Survey (CRS) — Student Version and Teacher Version were developed to address these limitations. In contrast with prior measures of student-teacher caring relationships, the CRS has a strong theoretical foundation based on Nel Noddings's conception of care that hypothesizes four interconnected aspects of care: Engrossment, Motivational Displacement, Reciprocity, and Attribution of Best Motive Consonant with Reality. It capitalizes on the combined strengths of prior measures of student-teacher relationships by incorporating and modifying relevant items from those surveys (Huffman, 2005; King & Chan, 2011; Tosolt, 2008), while simultaneously addressing limitations of these measures by including additional items written specifically for the CRS to more fully encompass the complexity and comprehensive nature of care in Noddings's definition. In addition, we directly acknowledged the relational aspect of care by developing parallel teacher and student versions of the survey. Our investigation of the psychometric properties of the CRS presented in this article provides support for this theoretical framework as well as suggesting potential ways that the theory could be expanded.

An important strength of this investigation involved randomly dividing the overall sample in half, with one sub-sample used for exploratory work (EFA) and the second used to confirm the factor structure (CFA). The results of the EFA provided some support for the four originally hypothesized item composites while also suggesting modifications to item groupings. Importantly, the EFA revealed an intriguing possibility of a fifth sub-construct connected with but not explicitly part of Noddings's definition of care. The items in this composite do not necessarily reflect traditionally recognized roles of teachers and therefore the new sub-construct was labeled Non-Traditional Roles. The items within this composite directly challenge teachers to consider their perceptions of what it means to be a caring teacher, specifically the tension between what may be seen as competing goals of attending to students' social and emotional needs versus their cognitive and academic growth.

The best-fitting CFA models for both students and teachers, fitted to the second sub-sample of data, include the new Non-Traditional Roles sub-construct, indicating its relevance to both teachers and students. The CFA also indicated subtle differences in the

best fitting models for teachers and students. For teachers a model with five first-order factors was slightly preferable, while for students the best-fitting model included a second order factor with the five first-order factors. Hence, the final model differs between students and teachers, leading to somewhat different interpretations of caring relationships for students and teachers. For students, caring relationships with their teachers appear to include one overarching construct of care that also includes five sub-constructs. In contrast, for teachers, caring relationships with their students involve more distinct components (the five sub-constructs). Generally speaking, these results suggest that students may have a more cohesive perspective of care than teachers.

Limitations and Implications

This study directly addresses the nature of relationships between students and their teachers, an issue relevant to a wide audience of practitioners and researchers interested in enhancing social-emotional learning and development in schools. The operationalization of such constructs as student-teacher caring relationships provides a way to develop quantitative measures to study potential links between teachers' and students' perceptions of care and students' personal growth, moral development, motivation to learn, and academic outcomes. Thus, the CRS could be a potent new tool for school counselors and practitioners interested in exploring ways to enhance climates of care in their schools and in promoting the development of reciprocal student-teacher relationships. Additionally, the subscale scores could be used to identify potential areas of growth in the different dimensions of care defined by Noddings.

However, additional research is necessary before the CRS is more widely used for policy and practice in educational settings. We turn to limitations of the current study and consider potential research to address them. First, we acknowledge the lack of racial and ethnic diversity in this sample as a limitation of the study. The construct of care could certainly differ depending on the race/ethnicity, culture, and socioeconomic background of teachers and students. The sample used for this study primarily consisted of white students and teachers from northern New England, so there are natural questions about the external validity of the survey results and underlying factor structures. Future research using more diverse samples from different cultural contexts will help determine whether the psychometric properties of the instruments vary by various demographic background characteristics of students and teachers.

A second limitation is connected with one of the study's major strengths, the development of parallel forms of the survey for teachers and students. In the current study, the teachers were not the teachers of the students who were surveyed, raising questions about whether conclusions can be drawn about the relational aspect of care, a central part of Noddings's theory. Future research could expand this investigation by having teachers fill out surveys for specific students, and vice versa, and then analyses of the student-teacher dyads could be conducted.

Third, although this study surveyed K–12 teachers and results can be used to determine whether perspectives on care vary by grade level, we do not know whether the factor structure itself differs for elementary school and secondary school teachers. Elementary and secondary classrooms differ substantially with respect to goals, focus, objectives, behavior management, and lesson plans. Such differences naturally impact teachers' notions of their roles and how they prioritize their tasks. In addition, students' different developmental stages across this wide grade spectrum impacts teachers' roles as care providers. An important next step in this research would involve investigating potential differences in factor structure based on student grade level.

Finally, the intriguing finding of the Non-Traditional Roles sub-construct confirmed in both student and teacher models needs additional investigation as well. It is possible that this sub-construct can provide a way for practitioners and researchers interested in student-teacher relationships to consider the evolving roles of teachers in contemporary education. Additional research with more varied samples, as described above, will help further explore whether this sub-construct is relevant across populations.

Acknowledgments

We gratefully acknowledge Dr. Barbara Houston for carefully reviewing different versions of the survey instruments as they were being developed and for providing constructive comments and critique at all stages of this research.

References

Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures*, *12*(1), 25–44. https://doi.org/10.1080/19312458.2017.1396583

Cothran, D. J., & Ennis, C. D. (2000). Building bridges to student engagement: Communicating respect and care for students in urban high schools. *Journal of Research and Development in Education*, *33*(2), 106–117.

DeVellis, R. F. (2003). *Scale development: Theory and application*. Sage.

Diller, A. (1988). The ethics of care and education: A new paradigm, its critics, and its educational significance. *Curriculum Inquiry*, *18*(3), 325–342. https://doi.org/10.2307/1179833

Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis: Understanding statistics*. Oxford University Press.

Gehlbach, H., & Brinkworth, M. E. (2011). Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology*, *15*(4), 380–387. https://doi.org/10.1037/a0025704

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Harvard University Press.

Goldstein, L. S., & Lake, V. E. (2000). "Love, love, and more love for children": Exploring preservice teachers' understandings of caring. *Teaching and Teacher Education*, *16*(8), 861–872. https://doi.org/10.1016/S0742-051X(00)00031-7

Goldstein, L. S., & Lake, V. E. (2003). The impact of field experience on preservice teachers' understandings of caring. *Teacher Education Quarterly*, *30*(3), 115–132. https://www.teqjournal.org/backvols/2003/30_3/goldstein&lake.pdf

Graham, A., Powell, M. A., & Truscott, J. (2016). Facilitating student well-being: Relationships do matter. *Educational Research*, *58*(4), 366–383. https://doi.org/10.1080/00131881.2016.1228841

Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, *72*(2), 625–638. https://doi.org/10.1111/1467-8624.00301

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huffman, V. L. (2005). *An instrument to measure the components of ethical caring as indicators of student teachers' proclivity to care* (Publication No. 3196649) [Doctoral dissertation, Illinois State University]. ProQuest Dissertations and Theses Global.

Hughes, J. N., Cavell, T. A., & Willson, V. (2001). Further support for the developmental significance of the quality of the teacher-student relationship. *Journal of School Psychology*, *39*(4), 289–301. https://doi.org/10.1016/S0022-4405(01)00074-7

Kelloway, E. K. (2015). *Using Mplus for structural equation modeling* (2nd ed.). Sage.

Kemp, H., & Reupert, A. (2012). "There's no big book on how to care": Primary pre-service teachers' experiences of caring. *Australian Journal of Teacher Education*, *37*(9), Article 7. https://doi.org/10.14221/ajte.2012v37n9.5

King, P. C., & Chan, T. C. (2011, October). *Teachers' and students' perceptions on teachers' caring behaviors*. Paper presented at GERA-36th annual meeting, Savannah, Georgia.

Lee, O., & Ravizza, D. (2008). Physical education pre-service teachers' conceptions of caring. *Education*, *128*(3), 460–472.

Lewis, J. L., Ream, R. K., Bocian, K. M., Cardullo, R. A., Hammond, K., & Fast, L. A. (2012). Con cariño: Teacher caring, math self-efficacy, and math achievement among Hispanic English learners. *Teachers College Record*, *114*(7), 1–42. https://doi.org/10.1177/016146811211400701

Mayeroff, M. (1971). *On caring*. Harper & Row.

Meyers, S. A. (2009). Do your students care whether you care about them? *College Teaching*, *57*(4), 205–210. https://doi.org/10.1080/87567550903218620

Muthén, L. K., & Muthén, B. O. (2017). *Mplus: Statistical analysis with latent variables*. *User's guide (Version 8)*. https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf

Newton, M., Fry, M., Watson, D., Gano-Overway, L., Kim, M., Magyar, M., & Guivernau, M. (2007). Psychometric properties of the Caring Climate Scale in a physical activity setting. *Revista de Psicología del Deporte*, *16*(1), 67–84.

Ng, P., Su, X. S., Chan, V., Leung, H., Cheung, W., & Tsun, A. (2013). The reliability and validity of a campus caring instrument developed for undergraduate students in Hong Kong. *Measurement and Evaluation in Counseling and Development*, *46*(2), 88–100. https://doi.org/10.1177/0748175612467463

Noddings, N. (1984). *Caring: A feminine approach to ethics and moral education*. University of California Press.

Noddings, N. (2003). *Caring: A feminine approach to ethics and moral education* (2nd ed.). University of California Press.

Phillippo, K. L., Conner, J., Davidson, S., & Pope, D. (2017). A systematic review of student self-report instruments that assess student-teacher relationships. *Teachers College Record*, *119*(8), 1–42. https://doi.org/10.1177/016146811711900801

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 365–386). Springer.

Robson, J., & Bailey, B. (2009). "Bowing from the heart": An investigation into discourses of professionalism and the work of caring for students in further education. *British Educational Research Journal*, *35*(1), 99–117. https://doi.org/10.1080/01411920802041731

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reportig structural equation modeling and confirmatory factor analysis results : A review. *The Journal of Educational Research*, *99*(6), 323–338. https://doi.org/10.3200/JOER.99.6.323-338

Teven, J. J., & McCroskey, J. C. (1997). The relationship of perceived teacher caring with student learning and teacher evaluation. *Communication Education*, *46*(1), 1–9. https://doi.org/10.1080/03634529709379069

Tomarken, A. J., & Waller, N. G. (2005). Structural equation modeling: Strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology*, *1*, 31–65. https://doi.org/10.1146/annurev.clinpsy.1.102803.144239

Tosolt, B. (2008). *Middle school students' perceptions of caring teacher behaviors: An empirical examination by student minority status* (Publication No. 3333077) [Doctoral dissertation, Oakland University]. ProQuest Dissertations and Theses Global.

Velasquez, A., West, R., Graham, C., & Osguthorpe, R. (2013). Developing caring relationships in schools: A review of the research on caring and nurturing pedagogies. *Review of Education*, *1*(2), 162–190. https://doi.org/10.1002/rev3.3014

Vogt, F. (2002). A caring teacher: Explorations into primary school teachers' professional identity and ethic of care. *Gender and Education*, *14*(3), 251–264. https://doi.org/10.1080/0954025022000010712

Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology*, *89*(3), 411–419. https://doi.org/10.1037/0022-0663.89.3.411

Williams, B., Brown, T., & Onsman, A. (2010). Exploratory factor analysis : A five-step guide for novices. *Journal of Emergency Primary Health Care*, *8*(3), Article 990399.

Williford, A. P., & Wolcott, C. S. (2015). SEL and student-teacher relationships. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice* (pp. 229–243). The Guilford Press.

Appendix

Path Diagrams for the Four Hypothesized CFA Models

(Models are the same for teachers and students, except that item #43 is not included in the student survey.)

Figure A1: Model 1. One Second-order Factor (Care) With Five First-order Factors Representing the Sub-constructs

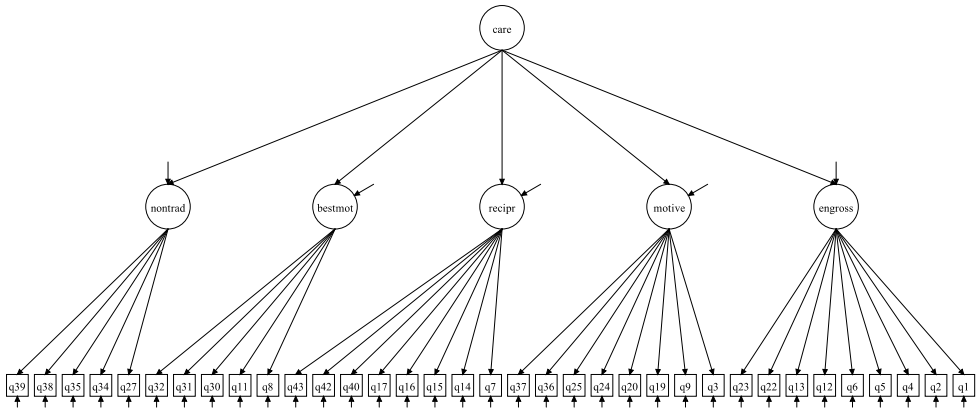


Figure A2: Model 2. Five First-order Factors With No Second-order Factor

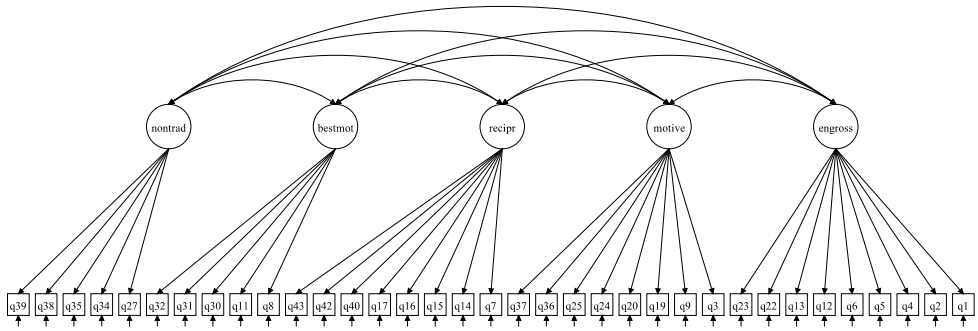


Figure A3: Model 3. One Second-order factor With Four First-order Factors

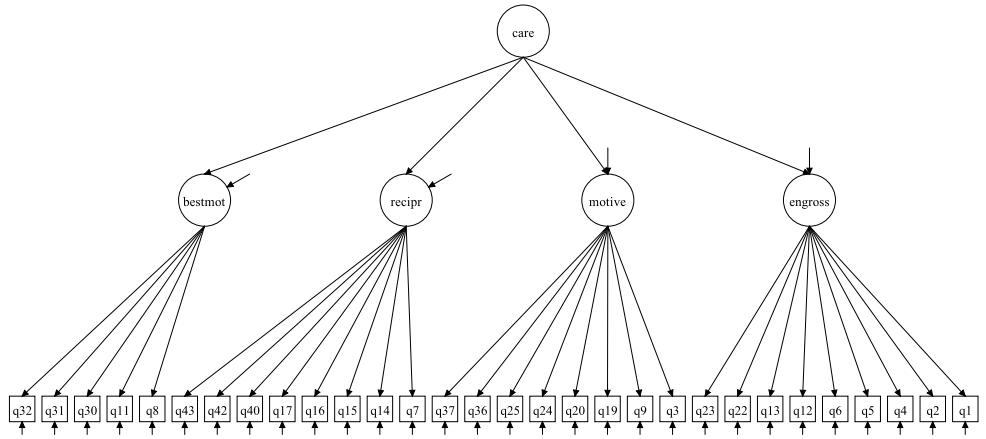
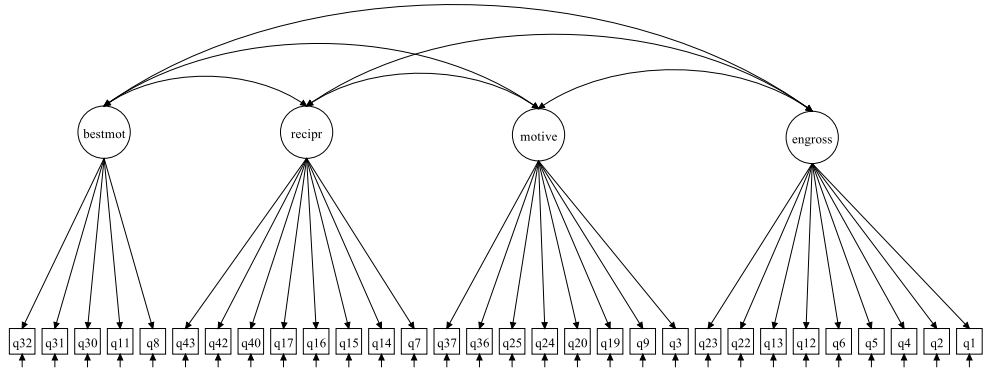


Figure A4: Model 4. Four First-order Factors With No Second-order Factor



以 Nel Noddings 關懷倫理學為基礎之師生關懷關係量表發展研究

張德忻、Suzanne E. Graham

摘 要

師生關懷關係對於學生的社交情緒和認知技能發展有着舉足輕重的角色，但用於了解師生關懷關係的量化研究經常受限於理論框架的表述不完整，導致關懷關係淪為空泛的詞彙。本研究的主要目的是開發根據完整理論背景並合乎心理計量學檢視的師生關懷關係量表。依據 Nel Noddings 關懷倫理學中的關懷現象，完整界定了校園內的師生關懷關係，其四個子構念包含：全神貫注、動機置換、互惠關係、歸因於與現實相輔相成的最佳動機。本研究共收回 772 名學生和 629 名老師的問卷，把老師和學生的問卷個別隨機分為兩組子樣本進行分析。使用探索性因子分析進行題目的刪減並了解項目組成，而後使用驗證性因子分析檢視四個假設的測量模型。結果顯示，與 Nel Noddings 的觀點相似，關懷關係包含多個維度。子構念的 Cronbach's alphas 介於 .68–.83 之間，視為可接受的信度範圍。此量表提供了一個以理論背景為基礎並同時考量學生和教師觀點的量表。後續需於更多不同的文化和區域施測，以進一步了解此問卷的外部效度。

關鍵詞：量表發展；關懷倫理學；師生關係；社交情緒學習

CHANG, Te-Hsin（張德忻）is Assistant Professor in the Holistic Education Center, Tzu Chi University of Science and Technology.

GRAHAM, Suzanne E. is Associate Professor in the Department of Education, University of New Hampshire.