Validation of a short form of the German form of the classroom community scale (CCS-D-SF)

Abstract

Background: Collaborative learning in medical education relies on both social integration and cognitive engagement. The classroom community scale (CCS) measures students' sense of community through two dimensions: *Connectedness* and *learning*. In German-speaking countries, the CCS has previously been validated in its full form (CCS-D). While an English short form (CCS-SF) exists, a validated German short form for efficient assessment was lacking. This study aimed to develop and validate a short form of the German CCS (CCS-D-SF).

Methods: The CCS-D was systematically reduced to a ten-item short form using principal component analysis (PCA) and iterative reliability analysis. Data on the short form were collected from N=171 first-semester medical students at the University of Tuebingen. Internal consistency was evaluated with Cronbach's alpha. Construct validity was assessed using confirmatory factor analysis (CFA) and PCA.

Results: The CCS-D-SF retained a two-factor structure corresponding to *connectedness* (i=5) and *learning* (i=5). The total scale showed good internal consistency (α=.81), with reliability for the subscales *connectedness* (α=.86) and *learning* (α=.65). CFA indicated an acceptable model fit (χ^2/df=2.16, RMSEA=.083, SRMR=.059, CFI=.93, TLI=.91), and PCA displayed also the factor structure.

Discussion: The CCS-D-SF is a valid, and practical instrument for assessing sense of community among German-speaking first-year medical students. Reliability ranges from marginal for *learning* to very good for *connectedness*. The overall reliability can be considered to be good. Questionnaires brevity reduces respondent burden without compromising psychometric quality and makes it suitable for educational research and classroom diagnostics. The CCS-D-SF may also inform interventions aiming to strengthen both social integration and learning engagement in collaborative medical education settings.

Keywords: medical education, interdisciplinary placement, confirmatory factor analysis, validation, classroom community

Harald Knof[1,2]
Victoria Huber[1,2]
Thomas Shiozawa[1,2]

1 Eberhard Karls University of Tuebingen, Faculty of Medicine, Department of Anatomy, Institute of Clinical Anatomy and Cell Analysis, Tuebingen, Germany

2 Eberhard Karls University of Tuebingen, Faculty of Medicine, TIME – Tuebingen Institute for Medical Education, Tuebingen, Germany

1. Introduction

Collaborative learning has increasingly gained importance in medical education as both a teaching and learning format. Such approaches promote teamwork, communication skills, and responsibility for individual as well as group performance - competencies that are essential for professional practice and the quality of healthcare delivery [1], [2]. In the following, the concepts of collaborative learning and sense of community are first elaborated, as they represent key theoretical foundations of the classroom community scale, a widely used instrument for measuring the sense of community in learning environments [3]. A more detailed conceptual framework has previously been provided by Knof and Shiozawa [4]. Subsequently, the rationale for the development of a short form of this scale is presented.

1.1. Collaborative learning in medical education

Collaborative and team-based learning approaches are increasingly used in higher education and medical education and can be implemented face-to-face, synchronously or asynchronously online [5], [6], [7].

Despite its benefits in fostering teamwork and communication, collaborative learning presents challenges. Participation may be unequal, communication ineffective, and interpersonal conflicts can arise, which can limit learning outcomes [8], [9], [10]. One key factor supporting effective collaboration is the degree to which students feel integrated into their learning environment. This construct, referred to as sense of community, is significantly associated with engagement, collaboration, and academic achievement in both face-to-face and online settings [11],

[12], [13], [14], [15]. Weak classroom community is linked to disengagement and dropout risk, highlighting the importance of fostering trust, collaboration, and shared commitment [16], [17], [18].

1.2. Sense of community

A widely cited concept of sense of community by McMillan and Chavis describes it as "a feeling that members have of belonging, that members matter to one another and to the group, and a shared faith that members' needs will be met through their commitment to be together" ([19], p.9). Essential elements include trust, mutual support, interactivity, shared values, and group cohesion [20], [21], [22], while the specific manifestation of this construct may vary depending on the context in which it occurs [23], [24], [25].

Classroom community comprises two main components: *Connectedness* measures feelings of interpersonal connection, while *learning* assesses perceptions of interaction and engagement within the classroom [21].

To measure classroom community, Rovai developed the Classroom Community Scale (CCS) based on this framework [3].

1.3. Classroom community scale

The classroom community scale (CCS) consists of 20 items across two subscales: *Connectedness* (10 odd-numbered items) and *learning* (10 even-numbered items) (see attachment 1, table S1) [3]. Participants respond on a 5-point Likert scale from *strongly disagree* to *strongly agree*. Total scores range from 20 to 80, with each subscale ranging from 10 to 40, where higher scores indicate a stronger sense of classroom community [3].

The CCS has been widely assessed in various countries across face-to-face, blended, and online learning environments [26], [27], [28], [29], [30], [31]. However, no validated German version was previously available. The authors translated the CCS into German (CCS-D) and found its validity comparable to the original and other translations [4].

1.4. Rationale for developing a short form

While long questionnaires in assessments offer certain advantages, they also present additional challenges for researchers [32]. Longer questionnaires are associated with increased respondent burden, which may contribute to participant fatigue, reduced response effort, incomplete responses, and higher refusal rates [33], [34]. Furthermore, longer instruments are often associated with higher refusal rates compared to shorter surveys [33]. These issues can negatively affect data quality and may ultimately threaten the validity of research findings [35], [36]. Consequently, the development of short forms is valuable not only because it reduces the burden placed on both participants and researchers, but also because shorter questionnaires can enhance data quality by lowering the number of missing responses, reducing aberrant response patterns, and increasing response rates [37]. The demand for shorter yet psychometrically robust scales have been widely acknowledged in survey and organizational research [33], [38], [39].

The validation of the long form of the CCS revealed cross-loadings of several items across the proposed dimensions [4]. These findings are consistent with previous reports showing substantial cross-loadings and challenges in establishing strong construct validity across different CCS versions [26], [31], [40], [41].

These cross-loadings indicate that items are not exclusively linked to a single latent dimension but instead capture elements of both *connectedness* and *learning*. Such patterns tend to weaken discriminant validity and may reduce the conceptual clarity of underlying constructs [42]. As a result, items with substantial loadings on more than one factor can add additional measurement error and create uncertainty in the interpretation of scale scores [43]. If cross-loadings are ignored, the resulting factor structure may become less stable, increasing the risk of model misspecification and potentially leading to inaccurate or flawed theoretical interpretations [44].

These considerations suggest that further refinement of the measurement model may be necessary.

One approach to instrument refinement is the development of a short form of the measure.

Consistent with established principles of scale construction, a parsimonious and clearly structured factor solution is generally preferred [32], [45]. Accordingly, item retention should prioritize items that load clearly on a single intended factor, ensuring that each item represents its designated content domain as precisely as possible [45]. In line with this, exploratory factor analysis procedures typically guide the removal of items with weak primary loadings or substantial loadings on multiple factors, and items that consistently exhibit cross-loadings are therefore often considered unsuitable for retention [42], [43].

Following Barnard-Brak and Shiu's recommendation to modify factor structure through exploratory factor analyses and item-level diagnostics to develop the instrument, Cho and Demmans Epp applied this approach using data from online learning environments [26], [41]. This work led to the development of an English short form of the CCS (CCS-SF), which has subsequently been validated in the literature [41], [46]. However, this version was derived from different datasets and a distinct contextual setting, namely online learning environments.

Given that the validation of the CCS-D already indicated differences in item loadings compared to the original CCS [4], it appeared necessary to develop a short form based on the present dataset. Because translation processes may introduce subtle semantic shifts affecting item interpretation, empirical validation within the German context is particularly important.

However, scale shortening brings methodological challenges. No methodological consensus exists regarding the optimal number of items to retain in short-form scale

development [47], and excessive item reduction may compromise content validity and reliability if construct domains are insufficiently represented [48]. Therefore, item selection must follow careful and transparent procedures to minimize potential losses in measurement quality [37].

2. Methods

2.1. Protocol for systematic item reduction of the CCS-D questionnaire

Given the rationale and challenges described above, based on the already translated version of the instrument, a systematic procedure for item reduction was developed and implemented. The goal was to shorten the questionnaire to a total of ten items – five per subscale – without substantially compromising its psychometric quality, theoretical foundation, or the content coverage of the subscales.

The decision to create a ten-item short version was based on several methodological and conceptual considerations. Methodological research suggests that less than four items per factor may result in insufficient internal reliability [47], [49]. Moreover, scale development literature emphasizes that instruments should include enough items to adequately represent the construct domain and prevent construct underrepresentation [50]. Consequently, recommendations for scale reduction stress the importance of balancing brevity with adequate construct coverage and reliability [33], [35], [37], [51].

Under classical test theory assumptions, internal consistency tends to increase with the number of items, in particular when using Cronbach's alpha, if items measure the same underlying construct [52], [53], [54]. Retaining a slightly larger number of items in a ten-item version compared to a more shortened eight-item version therefore contributes to more stable measurement properties. Furthermore, a scale of this length allows for broader content coverage of the underlying construct. This is essential for construct validity as item reduction may lead to insufficient representation of relevant aspects [48], [50]. Particularly in multidimensional constructs such as classroom community, overly aggressive item reduction may narrow the conceptual breadth of the instrument and weaken its ability to adequately represent the underlying construct [48], [50].

In addition, the ten-item version represents a practical compromise between measurement precision and efficiency, as it substantially reduces participants' response time while preserving the psychometric quality of the instrument [51].

The systematic reduction process was carried out in a three-step procedure. The approach was based both on the iterative methodology of Cho and Demmans Epp in their validation of their CCS short form [41], and on the guidelines of Hinkin [45].

The custom reduction protocol was predefined prior to the refinement procedure to minimize researcher-induced selection bias. The complete CCS-D dataset from Knof & Shiozawa served as the starting point [4]. To examine the factor structure, a Principal Component Analysis (PCA) with varimax rotation was conducted, while the internal consistency of the scales was determined using Cronbach's alpha. Subsequently, an iterative reduction process was initiated following a predefined protocol. The objective of the procedure was construct optimization while preserving content coverage rather than maximizing internal consistency coefficients.

In each iteration, the procedure began with step 1: All items that exhibited cross-loadings (secondary loadings ≥.30) or primary loadings on a non-target factor were removed simultaneously. Primary loadings on a non-target factor were defined as items loading more strongly on a factor other than the theoretically expected dimension based on the original scale structure and the framework by Rovai described above. These items do not clearly represent a single latent construct and may therefore compromise the validity of the scale [43], [53]. Whenever one or more items were removed in step 1, a new PCA was immediately conducted and Cronbach's alpha recalculated, marking the start of a new iteration. If no items met the criteria for removal in step 1, step 2 was applied: In this step, the item whose deletion produced an improvement in internal consistency (Cronbach's alpha if item deleted) was removed. After any removal in step 2, a PCA was performed, Cronbach's alpha recalculated, and the procedure returned to step 1 for the next iteration. If neither step 1 nor step 2 resulted in any removal, step 3 was executed, simultaneously eliminating all items with the lowest loadings on their theoretically corresponding factors in order to preserve the overall construct validity of the scale and maintain adequate representation of the underlying theoretical dimensions [43], [53]. As in the previous steps, a PCA and recalculation of Cronbach's alpha were performed after each removal, and the process restarted at step 1. This iterative procedure continued until no further items could be removed according to the predefined criteria, resulting in a psychometrically optimized short form with equal representation of both subscales (five items per factor), ensuring structural symmetry of the measurement model of classroom community.

The validation of the short form was conducted using an independent sample, following best-practice recommendations [33], [39], [48], [55]. This step is important because item responses may be influenced by the surrounding survey context [33]. Therefore, validating the shortened instrument on the same dataset used for item selection could introduce bias.

2.2. Data collection

At the end of the summer semester 2025, first-semester medical students at the Faculty of Medicine, University of Tuebingen, were surveyed using a paper-based ques-

tionnaire. By this time, participants had completed all courses of their first semester in the medical study program. Recruitment took place in person, conducted by the first and second author. Prior to participation, all students provided written informed consent after being informed about the purpose of the study, potential benefits and risks, their rights and responsibilities, and the voluntary nature of participation. Data were collected in pseudonymized form, and all participants consented to the publication of anonymized data. Participants were informed that they could withdraw their consent at any time without any negative consequences. Ethical approval for the study was granted by the ethics committee of Eberhard Karls University Tübingen (approval number: 086/2022B02).

2.3. Data analysis

Analogous to the validation of the CCS-D, the same validation procedure was applied [4].
Descriptive statistical analyses were conducted to characterize the sample and to determine the scale and item properties of the CCS-D-SF. Internal consistency was evaluated using Cronbach's alpha for each subscale as well as for the total scale. Values of alpha ≥.70 were considered satisfactory [54].
Construct validity of the two-factor structure of the CCS-D-SF was examined using confirmatory factor analysis (CFA) with maximum likelihood (ML) as used estimation method.
In the CFA, the observed variables (five *connectedness* items and five *learning* items) were specified as indicators of their respective latent variables (*connectedness* and *learning*). Model fit was evaluated according to widely accepted international standards [56], [57], [58].
The following indices and thresholds were used:

- Normed Chi-square (χ^2/df): acceptable fit indicated by values <3 [59].
- Comparative fit index (CFI) and Tucker-Lewis index (TLI): good model fit indicated by values ≥.95 [60].
- Root mean square error of approximation (RMSEA): values between .05 and .08 indicate acceptable fit [61], [62], [63].
- Standardized root mean square residual (SRMR): acceptable fit indicated by values <.08 [60].

To further examine the robustness of the model and explore underlying data-driven patterns not captured by the CFA, a PCA was conducted. PCA was performed using varimax rotation with Kaiser normalization. Items with factor loadings of ≥|.30| on one or both dimensions were retained [59], [64]. This complementary approach followed recommendations emphasizing the combined use of CFA and exploratory techniques to ensure model robustness [64], [65], [66].
All data preparation and descriptive analyses were conducted using IBM SPSS Statistics 30 (IBM Corp., Armonk, NY). The CFA was performed using IBM SPSS AMOS 26 (IBM Corp., Armonk, NY).

3. Results

3.1. Sample

A total of *N*=178 first-semester medical students at the Faculty of Medicine Tuebingen were surveyed using a paper-based questionnaire. 205 students could have been included; this represents a response rate of 86.8%. To meet the conditions for use of AMOS in further analysis, all participants with any missing data for items of CCS-D-SF were excluded. This resulted in 171 respondents.
Study participants were 68.4% female, corresponding to the gender distribution of the admission. The age ranged from 18 to 29 years; the mean age was 20.43 years (SD 1.97). Most participants were medical students (87.1%) followed by dentistry students (12.9%).

3.2. Systematic item reduction of the CCS-D questionnaire

The validation of the stepwise item reduction for the short form was based on the PCA of the CCS-D dataset, which demonstrated good internal consistency (Cronbach's alpha=.87; see attachment 1, table S2) [4].
Based on this analysis, step 1 of the item reduction protocol was implemented. All items exhibiting substantial cross-loadings (items 9, 10, 14, and 16) were removed. Subsequently, a PCA was conducted, and Cronbach's alpha was recalculated to provide the basis for the next iteration. This resulted in the first iteration of the CCS-D-SF with a Cronbach's alpha=.84 (see attachment 1, table S3).
In iteration 1 of the CCS-D-SF, no items displayed cross-loadings or primary loadings on a non-target factor. Therefore, step 2 was conducted, during which item 6 was removed because its exclusion increased Cronbach's alpha. The subsequent PCA and reliability analysis produced CCS-D-SF iteration 2 (Cronbach's alpha=.84) (see attachment 1, table S4).
During CCS-D-SF iteration 2, item 5 and item 17 were removed because of cross-loadings. A subsequent PCA and recalculation of Cronbach's alpha produced CCS-D-SF iteration 3 (Cronbach's alpha=.82) (see attachment 1, table S5).
In iteration 3 of the CCS-D-SF, item 1 was removed because it loaded on more than one factor. The subsequent PCA and reliability analysis yielded CCS-D-SF iteration 4 (Cronbach's alpha=.80) (see attachment 1, table S6).
In iteration 4, no items displayed cross-loadings or primary loadings on a non-target factor. Step 2 did not yield improvements in Cronbach's alpha, so step 3 was conducted, leading to the removal of items 7 and 2, which had the lowest loadings on their respective factors. The resulting 10-item version contained five items per subscale. A final PCA and recalculation of Cronbach's alpha produced CCS-D-SF iteration 5 (Cronbach's alpha=.79) (see attachment 1, table S7).

Table 1: Final version of CCS-D-SF

1	I feel connected to others in this course.
2	I feel that it is hard to get help when I have a question.
3	I trust others in this course.
4	I feel uneasy exposing gaps in my understanding.
5	I feel that I can rely on others in this course.
6	I feel that this course results in only modest learning.
7	I feel that members of this course depend on me.
8	I feel that my educational needs are not being met.
9	I feel confident that others will support me.
10	I feel that this course does not promote a desire to learn.

Table 2: Descriptive statistics of CCS-D-SF items

Descriptive statistics					
	N	Minimum	Maximum	Mean	Std. deviation
Item 1	171	0	4	3.15	.994
Item 2 i	171	0	4	2.84	1.002
Item 3	171	0	4	3.09	.900
Item 4 i	171	0	4	2.39	1.113
Item 5	171	0	4	3.09	.867
Item 6 i	171	0	4	2.64	.956
Item 7	171	0	4	2.70	.874
Item 8 i	171	0	4	2.92	.985
Item 9	171	0	4	3.04	.900
Item 10 i	171	0	4	2.81	.894
Valid N (listwise)	171				

Table 3: Descriptive statistics of CCS-D-SF subscales

Descriptive statistics					
	N	Minimum	Maximum	Mean	Std. deviation
CCS-D-SF-total	171	5.00	39.00	28.67	5.787
CCS-D-SF-connectedness	171	2.00	20.00	15.06	3.630
CCS-D-SF-learning	171	3.00	20.00	13.60	3.269
Valid N (listwise)	171				

To enhance usability, items were reordered to alternate between subscales, consistent with the original CCS format, and renumbered accordingly (see table 1). The final CCS-D-SF includes two subscales: *Connectedness* (five odd-numbered items) and *learning* (five even-numbered items). Each item is rated on a five-point Likert scale. According to the original CCS instrument for items 1, 3, 5, 7, 9, the scoring scale was: *strongly agree=4, agree=3, neutral=2, disagree=1, strongly disagree=0*; for items 2, 4, 6, 8, 10 the scoring scale was: *strongly agree=0, agree=1, neutral=2, disagree=3, strongly disagree=4* [3]. For further analysis these item-scores must be inverted. These items are labeled with an "i" in tables and figures of this paper.

3.3. Validation of the CCS-D-SF

3.3.1. Scale characteristics and internal consistency

Descriptive statistics for the individual items (see table 2) as well as the total CCS-D-SF and for each subscale are shown below (see table 3). The total CCS-D-SF (Cronbach's alpha=.81) and the subscale *connectedness* (alpha=.86) showed good internal consistency. The subscale *learning* is reported with a Cronbach's alpha=.65.

3.3.2. Confirmatory factor analysis

The CFA-Model for the 2-factor-model is displayed in figure 1.

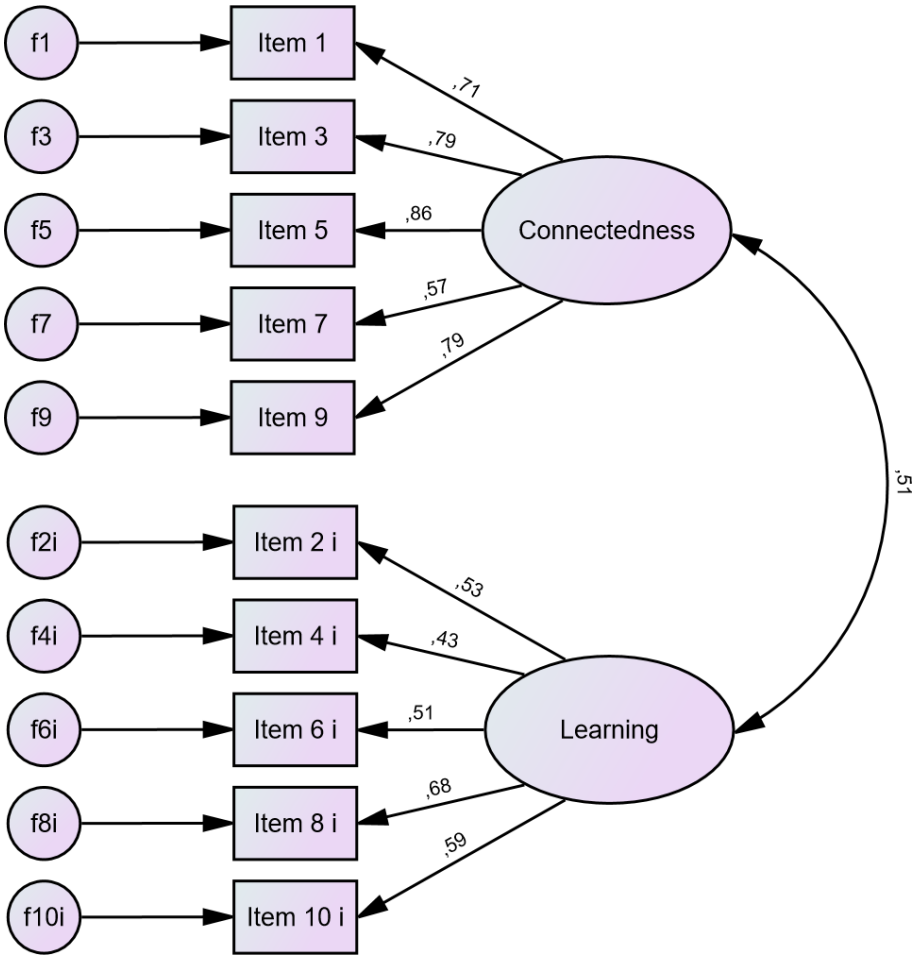


Figure 1: CCS-D-SF CFA Model

The two-factor CFA model demonstrated an overall acceptable model fit. The two latent factors were moderately correlated (r=.51). The normed chi-square (χ^2 [34, n=171]=73.529, p<.001; χ^2/df=2.163). indicated an adequate fit. The RMSEA (.083) exceeds the limit for an acceptable fit, while the SRMR (.059) indicated a good fit. The comparative fit indices, CFI (.93) and TLI (.91), were within the acceptable range but did not reach the threshold for a good fit.

3.3.3. Principal component analysis

The Kaiser-Meyer-Olkin measure was .811, indicating a relatively good suitability for factor analysis. Bartlett's test of sphericity was significant (*approx.* χ^2=581.215; *p*<.001), suggesting that the data do not form an identity matrix and are appropriate for factor analysis. Only factors with eigenvalues ≥1 were retained [68], [69]. Based on Kaiser's criteria and the scree plot, two factors with eigenvalues greater than 1 were retained, accounting for 54.9% of the total variance.

In the two-factor model rotated using varimax (see table 4), most items loaded primarily on a single factor, except for item 10 ("I feel that this course does not promote a desire to learn"). The underlying factor structure of the CCS-SF can thus be confirmed, with odd-numbered items loading highly on factor 1 and even-numbered items on factor 2, representing the dimensions of *connectedness* and *learning*.

4. Discussion

The CCS-D-SF shows good reliability and replicates the factor structure of the original version in the CFA. In a final PCA it could be ensured that the factor solution did not change after deleting items [51].

When considering model-fit indices, the CCS-D-SF performs better than the CCS-D.

A moderate correlation was found between the *connectedness* and *learning* subscales, suggesting that they share a substantial amount of variance while still representing distinct constructs, although the relationship indicates that the two dimensions of classroom community are not entirely independent. The marginally elevated RMSEA may partly reflect the substantial overlap between *connectedness* and *learning*. A likewise high correlation between the subscales has already been reported in the literature [30].

This relationship aligns with the *community of inquiry* framework, which posits that social presence *(connectedness)* and cognitive presence (*learning* engagement) interact dynamically to foster meaningful educational experiences [12], [70]. Students who feel socially integrated

Table 4: Varimax rotation PCA

Rotated component matrix[a]		
	Component	
	1	2
Item 1	.783	
Item 2 i		.565
Item 3	.816	
Item 4 i		.645
Item 5	.862	
Item 6 i		.661
Item 7	.637	
Item 8 i		.775
Item 9	.812	
Item 10 i	.343	.545

Extraction method: Principal component analysis.
Rotation method: Varimax with Kaiser normalization
a. Rotation converged in 3 iterations

within a course are more likely to engage cognitively, while deep cognitive engagement can, in turn, reinforce a sense of belonging and social connection [11]. The subscale *connectedness* showed very good internal consistency, whereas *learning* exhibited only marginally acceptable reliability. Lower reliability of the *learning* subscale is consistent with previous studies using the CCS and its derivatives, which have frequently reported lower internal consistency for learning-related items [26], [27], [71]. The reduced number of items in the CCS-D-SF may not fully capture this conceptual breadth, thereby lowering internal consistency and potentially affecting the correlation between subscales. Additionally, *learning* may be more context-dependent than connectedness, as students' perceptions of learning are influenced by instructional design, teaching style, and assessment practices.

The cross-loading of Item 10 ("I feel that this course does not promote a desire to learn") on the *connectedness* subscale further supports this interpretation. This pattern may reflect the way students perceive the term "course", which could be understood either as the content and structure of the class, the social group of participants, or a combination of both. Additionally, external factors such as the lecturer's personality and disposition can shape students' perceptions of social presence and classroom community, potentially affecting both *connectedness* and *learning* [72], [73]. Students' expectations about interactions with peers in future courses may also influence their motivation to develop a sense of community, thereby contributing to the observed subscale correlation [74].

From a pedagogical perspective, the interdependence between *connectedness* and *learning* carries important implications. Teaching strategies that intentionally foster social cohesion – such as group projects, peer mentoring, or structured reflection sessions – may not only strengthen connectedness but also enhance learning outcomes [12], [24]. Similarly, curriculum designs that emphasize collaborative knowledge construction can help to reinforce both social belonging and cognitive engagement. The CCS-D-SF could thus serve as a diagnostic instrument for educators to identify imbalances between these two facets and guide targeted interventions aimed at improving classroom community.

There are several limitations to this study that should be considered. A homogenization of student communities can be observed, driven by similar social and personal challenges, including uncertainties associated with the transition to higher education [75]. Typically, first-semester students focus on building personal relationships and establishing a professional network relevant to their studies [76]. This may influence how students perceive or prioritize *connectedness* and *learning*, as social integration and network-building could temporarily outweigh academic engagement or learning-focused interactions. In contrast to the present study, in which the sense of community was measured across an entire academic program, most previous research using the CCS has focused on individual courses or classroom sessions [77].

When interpreting the results of the scale reduction process, several methodological considerations should be acknowledged.

Factor analytic approaches are commonly used in scale shortening because factor loadings indicate the strength of the relationship between items and the underlying latent dimensions [33], [78]. Accordingly, many reduction procedures rely on psychometric indicators such as factor loadings, item-total correlations, and reliability estimates [35]. In the present study, scale refinement was not based solely on factor loadings; the internal consistency of the overall instrument was considered as a complementary criterion to ensure stability of the measurement model throughout the reduction process [50], [52]. However, reliance entirely on statistical criteria may produce shortened scales that are psychometrically homogeneous but conceptually narrow and instruments that are overly restrictive in content [34], [50], or instruments that cap-

ture slightly different aspects of the construct than the long form [33]. To address these concerns, attention was given to preserving an equal representation of the two underlying dimensions of classroom community, *connectedness* and *learning* [3].

The authors acknowledge that the present approach represents one possible strategy for scale reduction among several currently discussed methodological approaches. Short-form instruments have been commonly developed by iteratively removing items from existing longer questionnaires based on psychometric indicators [32]. Although these approaches are relatively simple to implement, they often consider only a limited range of psychometric criteria and may therefore result in shortcomings such as inadequate factor structures, insufficient reliability, or lack of measurement invariance [79]. More recently, algorithmic approaches have been proposed and shown promising results in the development of short forms [79], [80]. However, heuristic optimization procedures are not guaranteed to identify a single optimal solution and may produce different item combinations across repeated analyses [32]. These developments illustrate that scale shortening should not be understood as a purely mechanical optimization procedure, but rather as a context-dependent process requiring both statistical and theoretical considerations [79], [80].

As data-driven item reduction procedures may be sensitive to sample-specific characteristics, the resulting short form was evaluated using CFA in an independent validation sample [45], [55]. This approach also allows assessment of the stability and replicability of the factor structure, acknowledging that validity evidence obtained for long-form scales may not necessarily generalize to abbreviated versions [45], [55]. Alternative strategies such as split-sample validation within the original dataset could have reduced the risk of sample-specific overfitting. However, implementing a split-sample approach would have decreased the effective sample size available for item selection and factor analysis, which is particularly problematic, given that factor analysis can become unstable in smaller samples [63], [81]. The present study did not include external validation criteria that would allow a direct comparison of the validity of the short form. Without such external anchors, it cannot be fully determined whether this instrument captures the same underlying construct [33].

Improvements in model fit following scale reduction should not be interpreted as evidence of superior measurement quality per se, as fit indices are sensitive to model complexity, instrument length, and item homogeneity [82], [83].

Future research should therefore examine the relationships between the CCS-D-SF and theoretically related constructs in order to further establish its construct validity.

5. Conclusion

This study presented the systematic development and validation of a short form of the German version of the classroom community scale (CCS-D-SF).

The CCS-D-SF demonstrated good reliability, comparable to previous findings in the literature. Confirmatory factor analysis confirmed the two-factor structure of the scale, with model-fit indices even slightly better than those reported for the long version of the CCS-D. The factor structure was also supported by the principal component analysis.

Overall, the CCS-D-SF provides a practical and psychometrically sound instrument for measuring the sense of community among first-year medical students. Its brevity allows for reduced participant burden while measurement quality was maintained or even improved.

Future research should focus on applying the CCS-D-SF across German-speaking countries and further exploring its utility in diverse educational contexts.

Conference presentation

Excerpts from this manuscript were presented at the Annual Conference of the DACH Association for Medical Education (GMA) in Dresden 2026 [84].

Authors' ORCIDs

- Harald Knof: [0000-0002-0942-653X]
- Thomas Shiozawa: [0000-0002-7112-1016]

Acknowledgements

The authors would like to thank Prof Alfred P. Rovai from Regent University (Virginia Beach, VA, USA) for the permission to use and translate his CCS instrument.

The authors would like to thank Prof. Dr. Peter Martus for statistical consulting.

The authors are grateful to students at the Faculty of Medicine Tuebingen for participating in this study.

The authors acknowledge support from the Open Access Publication Fund of the University of Tuebingen.

Competing interests

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this article. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Attachments

Available from https://doi.org/10.3205/zma001891

1. Attachment_1.pdf (154 KB)
   Supplementary tables

References

1. Rosen MA, DiazGranados D, Dietz AS, Benishek LE, Thompson D, Pronovost PJ, Weaver SJ. Teamwork in healthcare: Key discoveries enabling safer, high-quality care. Am Psychol. 2018;73(4):433-450. DOI: 10.1037/amp0000298
2. Van Leuven AJ, Szymik BG, Ramsey LM, Hesse DW. A multi-year evaluation of medical student performance on and perceptions of collaborative gross anatomy laboratory examinations. Anat Sci Educ. 2023;16(3):557-566. DOI: 10.1002/ase.2223
3. Rovai AP. Development of an instrument to measure classroom community. Internet High Educ. 2002;5(3):197-211. DOI: 10.1016/s1096-7516(02)00102-1
4. Knof H, Shiozawa T. Validation of the German form of Classroom-Community-Scale (CCS-D). GMS J Med Educ. 2025;42(1):Doc13. DOI: 10.3205/zma001737
5. Dillenbourg P. What do you mean by 'collaborative learning'? In: Dillenbourg P, editor. Collaborative Learning: Cognitive and Computational Approaches. Advances in learning and instruction series. 1st ed. London: Pergamon; 1999. (Advances in learning and instruction series). p.1-15.
6. Sevenhuysen S, Haines T, Kiegaldie D, Molloy E. Implementing collaborative and peer-assisted learning. Clin Teach. 2016;13(5):325-331. DOI: 10.1111/tct.12583
7. O'Brien BC, Forrest K, Wijnen□Meijer M, ten Cate O. A Global View of Structures and Trends in Medical Education. In: Swanwick T, Forrest K, O'Brien BC, editors. Understanding Medical Education. 3rd ed. Hoboken: Wiley-Blackwill; 2018. p.7-22. DOI: 10.1002/9781119373780.ch2
8. Almajed A, Winning T, Skinner V, Peterson R. Medical and Dental Students' Perceptions and Experiences of Collaborative Learning: A Systematic Review. JBI Libr Syst Rev. 2011;9(32 Suppl):1-13. DOI: 10.11124/01938924-201109321-00001
9. Skinner VJ, Braunack-Mayer A, Winning TA. The Purpose and Value for Students of PBL Groups for Learning. Interdiscip J Problem-Based Learn. 2015;9(1). DOI: 10.7771/1541-5015.1499
10. Pluta WJ, Richards BF, Mutnick A. PBL and beyond: trends in collaborative learning. Teach Learn Med. 2013;25 Suppl 1:S9-S16. DOI: 10.1080/10401334.2013.842917
11. Chatterjee R, Correia A-P. Online Students' Attitudes Toward Collaborative Learning and Sense of Community. Am J Distance Educ. 2019;34(1):53-68. DOI: 10.1080/08923647.2020.1703479
12. Garrison DR, Anderson T, Archer W. The first decade of the community of inquiry framework: A retrospective. Internet High Educ. 2010;13(1-2):5-9. DOI: 10.1016/j.iheduc.2009.10.003
13. Halic O, Lee D, Paulus T, Spence M. To blog or not to blog: Student perceptions of blog effectiveness for learning in a college-level course. Internet High Educ. 2010;13(4):206-213. DOI: 10.1016/j.iheduc.2010.04.001
14. Lai KW. Knowledge construction in online learning communities: a case study of a doctoral course. Stud High Educ. 2013;40(4):561-579. DOI: 10.1080/03075079.2013.831402
15. Rovai AP. Building Sense of Community at a Distance. Int Rev Res Open Dis. 2002;3(1):1-16. DOI: 10.19173/irrodl.v3i1.79
16. Bruffee KA. Collaborative Learning: Higher Education, Interdependence, and the Authority of Knowledge. 1st ed. Baltimore: Johns Hopkins University Press; 1993.
17. Tinto V. Classrooms as Communities. J High Educ. 1997;68(6):599-623. DOI: 10.1080/00221546.1997.11779003
18. Sheets MF. Characteristics of Adult Education Students and Factors Which Determine Course Completion: A Review. New Horiz Adult Educ. 1992;6(1):3-18. DOI: 10.1002/nha3.10034
19. McMillan D, Chavis D. Sense of Community: A Definition and Theory. J Community Psychol. 1986;14(1):6-23. DOI: 10.1002/1520-6629(198601)14:1<6::AID-JCOP2290140103>3.0.CO;2-I
20. Hill JL. Psychological sense of community: Suggestions for future research. J Community Psychol. 1996;24(4):431-438. DOI: 10.1002/(Sici)1520-6629(199610)24:4<431::Aid-Jcop10>3.3.Co;2-Z
21. Rovai AP. Sense of community, perceived cognitive learning, and persistence in asynchronous learning networks. Internet High Educ. 2002;5(4):319-332. DOI: 10.1016/s1096-7516(02)00130-6
22. Shaffer C, Anundsen K. Creating community anywhere: Finding support and connection in a fragmented world. 1st ed. Los Angeles: Tarcher/Perigree; 1993.
23. Rheingold H. The virtual community: Finding commection in a computerized world. 1st ed. Boston: Addison-Wesley Longman Publishing Co. Inc.; 1993.
24. Strayhorn TL. College Students' Sense of Belonging - A Key to Educational Success for All Students. 2nd ed. New York: Routlege; 2018. DOI: 10.4324/9781315297293
25. Wang W, Gao Y, Pitts A, Dong L. A Bibliometric Analysis of Neighborhood Sense of Community. Sustainability. 2023;15(5):4183. DOI: 10.3390/su15054183
26. Barnard-Brak L, Shiu W. Classroom Community Scale in the blended learning environment: A psychometric review. Int J E-Learn. 2010;9(3):303-311.
27. Hur EH, Glassman M, Kim Y. Finding autonomy in activity: Development and validation of a democratic classroom survey. Educ Ass Eval Account. 2013;25(4):303-320. DOI: 10.1007/s11092-013-9173-y
28. Ahmady S, Kohan N, Bagherzadeh R, Rakshhani T, Shahabi M. Validity testing of classroom community scale in virtual environment learning: A cross sectional study. Ann Med Surg (Lond). 2018;36:256-260. DOI: 10.1016/j.amsu.2018.08.021
29. Yan J, Ding X, Xiong L, Liu E, Zhang Y, Luan Y, Qin L, Zhou C, Zhang W. Team-based learning: assessing the impact on anatomy teaching in People's Republic of China. Adv Med Educ Pract. 2018;9:589-594. DOI: 10.2147/AMEP.S169949
30. Abdeldayem M, Aldulaimi SH, Aldulaimi MLA. Virtual Learning and Students' Connectedness in the Time of Coronavirus. Int J Adv Sci. 2020;29(5):12634-12645.
31. Perrucci V, Cacciamani S, Khanlari A, Balboni G. The Classroom Community Scale in a Face-to-Face University Context: Factorial Structure and Convergent/Divergent Validity. Test Psychom Methodol Appl Psychol. 2022;29(3):297-307. DOI: 10.4473/TPM29.3.2
32. Franzen P, Arens AK, Greiff S, van der Westhuizen L, Fischbach A, Wollschläger R, Niepl C. Developing and Validating a Short-Form Questionnaire for the Assessment of Seven Facets of Conscientiousness in Large-Scale Assessments. J Pers Assess. 2022;104(6):759-773. DOI: 10.1080/00223891.2021.1998083

33. Stanton JM, Sinar EF, Balzer WK, Smith PC. Issues and Strategies for Reducing the Length of Self□Report Scales. Pers Psychol. 2002;55(1):167-194. DOI: 10.1111/j.1744-6570.2002.tb00108.x

34. Passarelli M, Casetta L, Rizzi L, Chiorri C, Cassina F, Voi S, Rocco D. Short and sweet: Comparing strategies for the reduction of questionnaires on self-criticism and social safeness while preserving construct validity. Int J Psychol. 2024;59(6):1234-1244. DOI: 10.1002/ijop.13249

35. Kilmen S, Bulut O. Shortening Psychological Scales: Semantic Similarity Matters. Educ Psychol Meas. 2025;85(5):910-934. DOI: 10.1177/00131644251319047

36. Yarkoni T. The Abbreviation of Personality, or how to Measure 200 Personality Scales with 200 Items. J Res Pers. 2010;44(2):180-198. DOI: 10.1016/j.jrp.2010.01.002

37. Alcaraz S, Viladrich C, Torregrosa M. Less time, better quality. Shortening questionnaires to assess team environment and goal orientation. Span J Psychol. 2013;16:E77. DOI: 10.1017/sjp.2013.74

38. Fraser BJ. Development of Short Forms of Several Classroom Environment Scales. J Educ Meas. 1982;19(3):221-227. DOI: 10.1111/j.1745-3984.1982.tb00130.x

39. Coste J, Guillemin F, Pouchot J, Fermanian J. Methodological approaches to shortening composite measurement scales. J Clin Epidemiol. 1997;50(3):247-252. DOI: 10.1016/s0895-4356(96)00363-0

40. Lewis K, McVay-Dyche J, Chen H, Soto T. Examining Sense of Community among Medical Professionals in an Online Graduate Program. J Educ Online. 2015;12(1). DOI: 10.9743/jeo.2015.1.8

41. Cho J, Demmans Epp C. Improving the Classroom Community Scale: Toward a Short-Form of the CCS. In: American Educational Research Association (AERA) Annual Meeting; 2019 Apr 5-9; Toronto, Canada. DOI: 10.3102/1432534

42. Brown TA. Confirmatory factor analysis for applied research. 2nd ed. New York (NY): The Guilford Press; 2015.

43. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis. 8th ed. Cheriton House: Cengage Learning; 2019.

44. Li Y, Wen Z, Hau KT, Yuan KH, Peng Y. Effects of Cross-loadings on Determining the Number of Factors to Retain. Struct Equation Model. 2020;27(6):841-863. DOI: 10.1080/10705511.2020.1745075

45. Hinkin TR. A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires. Organ Res Method. 1998;1(1):104-121. DOI: 10.1177/109442819800100106

46. Tackett M, Viel S, Manturuk K. A validation of the short-form classroom community scale for undergraduate mathematics and statistics students. J Uni Teach Learn Pract. 2023;20(1). DOI: 10.53761/1.20.01.08

47. Jokovic A, Locker D, Guyatt G. Short forms of the Child Perceptions Questionnaire for 11-14-year-old children (CPQ11-14): development and initial evaluation. Health Qual Life Outcomes. 2006;4:4. DOI: 10.1186/1477-7525-4-4

48. Smith GT, McCarthy DM, Anderson KG. On the sins of short-form development. Psychol Assess. 2000;12(1):102-111. DOI: 10.1037//1040-3590.12.1.102

49. Watson D, Clark LA. Measurement and mismeasurement of mood: recurrent and emergent issues. J Pers Assess. 1997;68(2):267-296. DOI: 10.1207/s15327752jpa6802_4

50. Clark LA, Watson D. Constructing validity: Basic issues in objective scale development. Psychol Assess. 1995;7(3):309-319. DOI: 10.1037/14805-012

51. Worthington RL, Whittaker TA. Scale Development Research: A Content Analysis and Recommendations for Best Practices. Couns Psychol. 2006;34(6):806-838. DOI: 10.1177/0011000006288127

52. Cortina JM. What is coefficient alpha? An examination of theory and applications. J Appl Psychol. 1993;78(1):98-104. DOI: 10.1037/0021-9010.78.1.98

53. Osborne J. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. Scotts Valley, CA: CreateSpace Independent Publishing; 2014.

54. Nunnally JC. Psychometric theory. 2nd ed. New York (NY): McGraw-Hill; 1978.

55. Kruyen PM, Emons WH, Sijtsma K. On the Shortcomings of Shortened Tests: A Literature Review. Int J Testing. 2013;13(3):223-248. DOI: 10.1080/15305058.2012.703734

56. Crowley SL, Fan X. Structural equation modeling: basic concepts and applications in personality assessment research. J Pers Assess. 1997;68(3):508-531. DOI: 10.1207/s15327752jpa6803_4

57. Hooper D, Coughlan J, Mullen M. Structural Equation Modeling: Guidelines for Determining Model Fit. Elect J Business Res Meth. 2007;6.

58. Kline RB. Principles and practice of structural equation modeling. 4th ed. New York (NY): Guilford Press; 2016.

59. Kline RB. Principles and practice of structural equation modeling. New York (NY): Guilford Press; 1998.

60. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equat Mod. 1999;6(1):1-55. DOI: 10.1080/10705519909540118

61. Browne MW, Cudeck R. Alternative Ways of Assessing Model Fit. Sociol Methods Res. 1992;21(2):230-258. DOI: 10.1177/0049124192021002005

62. Jöreskog KG, Sörbom D. LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language. Scientific Software International; 1993.

63. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods. 1996;1(2):130-149.

64. Knekta E, Runyon C, Eddy S. One Size Doesn't Fit All: Using Factor Analysis to Gather Validity Evidence When Using Surveys in Your Research. CBE Life Sci Educ. 2019;18(1):rm1. DOI: 10.1187/cbe.18-04-0064

65. Martin CR, Savage-McGlynn E. A 'good practice' guide for the reporting of design and analysis for psychometric evaluation. J Reprod Infant Psychol. 2013;31(5):449-455. DOI: 10.1080/02646838.2013.835036

66. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. Psychol Methods. 1999;4(3):272-299. DOI: 10.1037/1082-989x.4.3.272

67. Comrey AL, Lee HB. A first course in factor analysis. 2nd ed. Hillsdale (NJ): L. Erlbaum Associates; 1992.

68. Kaiser HF. The Application of Electronic Computers to Factor Analysis. Educ Psychol Meas. 1960;20(1):141-151. DOI: 10.1177/001316446002000116

69. Guttman L. Some Necessary Conditions for Common-Factor Analysis. Psychometrika. 2025;19(2):149-161. DOI: 10.1007/bf02289162

70. Arbaugh JB, Cleveland-Innes M, Diaz SR, Garrison DR, Ice P, Richardson JC, Swan KP. Developing a community of inquiry instrument: Testing a measure of the Community of Inquiry framework using a multi-institutional sample. Internet High Educ. 2008;11(3):133-136. DOI: 10.1016/j.iheduc.2008.06.003

71. Zhang YY, Lin X, Xu M. Rovai's Classroom Community Scale and Its Application in Chinese College English Class. Sino-US Engl Teach. 2011;8(9):592-598.

72. Lowenthal PR, Trespalacios J. Classroom Community and Time: Comparing Student Perceptions in Traditional vs. Accelerated Online Courses. Online Learn. 2022;26(4):59-77. DOI: 10.24059/olj.v26i4.3498

73. Trespalacios J, Lowenthal P. What do they really like? An investigation of students' perceptions of their coursework in a fully online educational technology program. Australas J Educ Technol. 2019;35(5):60-78. DOI: 10.14742/ajet.4364

74. Lowenthal PR, Dunlap JC. Investigating students' perceptions of instructional strategies to establish social presence. Distance Educ. 2018;39(3):281-298. DOI: 10.1080/01587919.2018.1476844

75. Picton A, Greenfield S, Parry J. Why do students struggle in their first year of medical school? A qualitative study of student voices. BMC Med Educ. 2022;22(1):100. DOI: 10.1186/s12909-022-03158-4

76. Traus A, Höffken K, Thomas S, Mangold K, Schröer W. Stu.diCo. – Studieren digital in Zeiten von Corona. Hildesheim: UVH - Universitätsverlag Hildesheim; 2020.

77. Beeson E, Aideyan B, Shoney CO, Bowes DA, Ansell KL, Peterson HM. Predicting Sense of Community among Graduate Students in a Distance Learning Environment. Univers J Educ Res. 2019;7(3):746-753. DOI: 10.13189/ujer.2019.070314

78. Floyd FJ, Widaman KF. Factor analysis in the development and refinement of clinical assessment instruments. Psychol Assess. 1995;7(3):286-299. DOI: 10.1037/1040-3590.7.3.286

79. Olaru G, Witthöft M, Wilhelm O. Methods matter: Testing competing models for designing short-scale Big-Five assessments. J Res Person. 2015;59:56-68. DOI: 10.1016/j.jrp.2015.09.001

80. Schroeders U, Wilhelm O, Olaru G. The influence of item sampling on sex differences in knowledge tests. Intelligence. 2016;58:22-32. DOI: 10.1016/j.intell.2016.06.003

81. MacCallum R, Widaman K, Zhang S, Hong S. Sample Size in Factor Analysis. Psychol Methods. 1999;4:84-99. DOI: 10.1037/1082-989X.4.1.84

82. Kenny DA, McCoach DB. Effect of the Number of Variables on Measures of Fit in Structural Equation Modeling. Struct Equ Modeling. 2009;10(3):333-351. DOI: 10.1207/s15328007sem1003_1

83. Jackson DL, Gillaspy JA, Purc-Stephenson R. Reporting practices in confirmatory factor analysis: an overview and some recommendations. Psychol Methods. 2009;14(1):6-23. DOI: 10.1037/a0014694

84. Knof H, Huber V, Shiozawa T. Validierung einer Kurzform der Deutschen Version der Classroom Community Scale (CCS-D-SF). In: Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). Dresden, 17.-19.09.2026. Düsseldorf: German Medical Science GMS Publishing House; 2026. DocV-0701. DOI: 10.3205/26gma037

Corresponding author:
Dr. med. Thomas Shiozawa
Eberhard Karls University of Tuebingen, Faculty of Medicine, Department of Anatomy, Institute of Clinical Anatomy and Cell Analysis, Elfriede-Aulhorn-Str. 8, D-72076 Tuebingen, Germany
thomas.shiozawa@uni-tuebingen.de

Please cite as
Knof H, Huber V, Shiozawa T. Validation of a short form of the German form of the classroom community scale (CCS-D-SF). GMS J Med Educ. 2026;43(7):Doc97.
DOI: 10.3205/zma001891, URN: urn:nbn:de:0183-zma0018918

This article is freely available from
https://doi.org/10.3205/zma001891

Received: *2025-11-01*
Revised: *2026-07-17*
Accepted: *2026-08-10*
Published: *2026-11-16*

Copyright
©2026 Knof et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at http://creativecommons.org/licenses/by/4.0/.

Validierung einer Kurzform der Deutschen Version der Classroom Community Scale (CCS-D-SF)

Harald Knof[1,2]
Victoria Huber[1,2]
Thomas Shiozawa[1,2]

1 Eberhard Karls Universität Tübingen, Medizinische Fakultät, Institut für Klinische Anatomie und Zellanalytik, Tübingen, Deutschland

2 Eberhard Karls Universität Tübingen, Medizinische Fakultät, TIME – Tübingen Institute for Medical Education, Tübingen, Deutschland

Zusammenfassung

Hintergrund: Kollaboratives Lernen in der medizinischen Ausbildung basiert sowohl auf sozialer Integration als auch auf kognitivem Engagement. Die Classroom Community Scale (CCS) erfasst das Gemeinschaftsgefühl von Studierenden anhand von zwei Dimensionen: *Verbundenheit* und *Lernen*. In deutschsprachigen Ländern wurde die CCS bereits in ihrer Langform (CCS-D) validiert. Während eine englischsprachige Kurzform (CCS-SF) existiert, fehlte bislang eine validierte deutsche. Ziel dieser Studie war es, eine Kurzform der deutschen CCS (CCS-D-SF) zu entwickeln und zu validieren.

Methoden: Die CCS-D wurde systematisch auf eine Kurzform mit zehn Items reduziert unter Verwendung von Hauptkomponentenanalysen (PCA) und iterativen Reliabilitätsanalysen. Die Daten für die Kurzform wurden von N=171 Medizinstudierenden im ersten Semester an der Universität Tübingen erhoben. Die interne Konsistenz wurde anhand von Cronbachs Alpha bewertet. Die Konstruktvalidität wurde mithilfe einer konfirmatorischen Faktorenanalyse (CFA) und einer PCA untersucht.

Ergebnisse: Die CCS-D-SF behielt eine Zwei-Faktoren-Struktur bei, die den Dimensionen *Verbundenheit* (5 Items) und *Lernen* (5 Items) entspricht. Die Gesamtskala zeigte eine gute interne Konsistenz (α=.81). Die Reliabilität der Subskalen *Verbundenheit* α=.86 und *Lernen* α=.65. Die CFA zeigte einen akzeptablen Modellfit (χ^2/df=2.16, RMSEA=.083, SRMR=.059, CFI=.93, TLI=.91). Auch die PCA bestätigte die Zwei-Faktoren-Struktur.

Diskussion: Die CCS-D-SF stellt ein valides und praktikables Instrument zur Erfassung des Gemeinschaftsgefühls bei Medizinstudierenden im ersten Studienjahr dar. Die Reliabilität reicht von einer knapp akzeptablen Reliabilität für *Lernen* bis zu einer sehr guten Reliabilität für *Verbundenheit.* Die Reliabilität der Gesamtskala kann als gut eingestuft werden. Die Kürze des Fragebogens reduziert die Belastung der Teilnehmenden, ohne die psychometrische Qualität wesentlich zu beeinträchtigen, und macht das Instrument für die Bildungsforschung sowie die diagnostische Erfassung der Klassengemeinschaft geeignet. Die CCS-D-SF kann zudem Hinweise für Interventionen liefern, die darauf abzielen, die soziale Integration sowie das Lernengagement in kollaborativen Lernumgebungen zu stärken.

Schlüsselwörter: medizinische Ausbildung, interdisziplinäre Verwendung, konfirmatorische Faktorenanalyse, Validierung, Classroom Community

1. Einleitung

Kollaboratives Lernen hat in der medizinischen Ausbildung zunehmend an Bedeutung gewonnen, sowohl als Lehr- als auch als Lernformat. Solche Ansätze fördern Teamarbeit, Kommunikationsfähigkeiten und die Übernahme von Verantwortung für die individuelle sowie die gemeinsame Leistung. Diese Kompetenzen sind für die professionelle Berufsausübung und die Qualität der Gesundheitsversorgung von entscheidender Bedeutung [1], [2]. Im Folgenden werden zunächst die Konzepte des kollaborativen Lernens und des Gemeinschaftsgefühls näher erläutert, da sie die zentralen theoretischen Grundlagen der Classroom Community Scale darstellen, einem weit verbreiteten Instrument zur Messung des Gemeinschaftsgefühls in Lernumgebungen [3]. Ein detaillierterer konzeptioneller Rahmen wurde bereits von Knof und Shiozawa vorgestellt [4]. Anschließend wird die Begründung für die Entwicklung einer Kurzform dieser Skala dargestellt.

1.1. Kollaboratives Lernen in der Medizinischen Ausbildung

Kollaborative und teambasierte Lernansätze werden zunehmend in der Hochschulbildung und der medizinischen Ausbildung eingesetzt und können in Präsenz, sowie synchron oder asynchron online umgesetzt werden [5], [6], [7].
Trotz seiner Vorteile für die Förderung von Teamarbeit und Kommunikation bringt kollaboratives Lernen auch Herausforderungen mit sich. Die Beteiligung kann ungleich verteilt sein, die Kommunikation kann ineffektiv sein und zwischenmenschliche Konflikte können entstehen, was Lernergebnisse beeinträchtigen kann [8], [9], [10]. Ein wesentlicher Faktor für eine effektive Zusammenarbeit ist der Grad, in dem sich Studierende in ihre Lernumgebung integriert fühlen. Dieses als Gemeinschaftsgefühl bezeichnete Konstrukt steht sowohl in Präsenz- als auch in Online-Lernumgebungen in einem signifikanten Zusammenhang mit Engagement, Zusammenarbeit und akademischen Leistungen [11], [12], [13], [14], [15]. Ein schwaches Gemeinschaftsgefühl innerhalb der Lerngruppe steht mit einer geringeren Beteiligung und einem erhöhten Abbruchrisiko in Zusammenhang und unterstreicht die Bedeutung der Förderung von Vertrauen, Zusammenarbeit und gemeinsamer Verantwortung [16], [17], [18].

1.2. Gemeinschaftsgefühl

Eine weit verbreitete Definition des „sense of community"(Gemeinschaftsgefühl) stammt von McMillan und Chavis: „sense of community ist das Gefühl, dass Mitglieder ein Gefühl der Zugehörigkeit haben, das Gefühl, dass die Mitglieder einander und der Gruppe wichtig sind, und ein gemeinsamer Glaube, dass die Bedürfnisse der Mitglieder durch ihr Engagement, zusammen zu sein, erfüllt werden" ([19], p.9). Zu den wesentlichen Elementen gehören Vertrauen, gegenseitige Unterstützung, Interaktivität, gemeinsame Werte und Gruppenzusammenhalt [20], [21], [22], wobei die konkrete Ausprägung dieses Konstrukts je nach Kontext, in dem es auftritt, variieren kann [23], [24], [25].
Eine Klassengemeinschaft als kollaborative Lernumgebung umfasst im Wesentlichen zwei Hauptkomponenten: *Verbundenheit* erfasst das Gefühl der zwischenmenschlichen Verbundenheit, während *Lernen* die Wahrnehmung von Interaktion und Beteiligung innerhalb der Lerngruppe erfasst [21]. Zur Messung des Gemeinschaftsgefühls in kollaborativen Lernumgebungen entwickelte Rovai die Classroom Community Scale (CCS) [3].

1.3. Classroom Community Scale

Die Classroom Community Scale (CCS) besteht aus 20 Items in zwei Subskalen: *Verbundenheit* (10 ungerade nummerierte Items) und *Lernen* (10 gerade-nummerierte Items) (siehe Anhang 1, Tabelle S1) [3]. Teilnehmende bewerten jedes Item auf einer 5-stufigen Likert-Skala von *„stimme überhaupt nicht zu"* bis *„stimme vollständig zu"*. Die Gesamtpunktzahl reicht von 20 bis 80, wobei jede Subskala Werte von 10 bis 40 umfasst. Höhere Werte weisen auf ein stärkeres Gemeinschaftsgefühl hin [3].
Die CCS wurde in verschiedenen Ländern und in unterschiedlichen Lernumgebungen, darunter Präsenzunterricht, Blended Learning und Online-Lernumgebungen, umfassend untersucht [26], [27], [28], [29], [30], [31]. Jedoch war bislang keine validierte deutschsprachige Version vorhanden. Die Autoren übersetzten die CCS ins Deutsche (CCS-D) und stellten fest, dass ihre Validität mit der der Originalversion und anderer Übersetzungen vergleichbar ist [4].

1.4. Rationale für die Entwicklung einer Kurzform

Während lange Fragebögen bei Erhebungen gewisse Vorteile bieten, bringen sie auch zusätzliche Herausforderungen für Forschende mit sich [32]. Längere Fragebögen gehen mit einer höheren Belastung der Teilnehmenden einher, was zu Ermüdung, einer geringeren Sorgfalt bei der Beantwortung und unvollständigen Antworten führen kann [33], [34]. Darüber hinaus weisen längere Instrumente im Vergleich zu kürzeren Befragungen häufig höhere Verweigerungsraten auf [33]. Diese Aspekte können sich negativ auf die Datenqualität auswirken und letztlich die Validität der Forschungsergebnisse beeinträchtigen [35], [36]. Die Entwicklung von Kurzformen ist daher nicht nur deshalb sinnvoll, weil sie die Belastung für Teilnehmende und Forschende reduziert, sondern auch, weil kürzere Fragebögen die Datenqualität verbessern können, indem sie die Anzahl fehlender Antworten verringern, auffällige Antwortmuster reduzieren und die Rücklaufquote erhöhen [37]. Der Bedarf an kürzeren und zugleich psychometrisch robusten Skalen ist in der Forschung weithin anerkannt [33], [38], [39].
Die Validierung der Langform der CCS zeigte Kreuzladungen mehrerer Items über die zugrundeliegenden theoretischen Dimensionen hinweg [4]. Diese Ergebnisse stimmen mit früheren Untersuchungen überein, die ebenfalls deutliche Kreuzladungen und Schwierigkeiten bei der Ermittlung einer aussagekräftigen Konstruktvalidität über verschiedene Versionen der CCS hinweg festgestellt haben [26], [31], [40], [41].
Diese Kreuzladungen weisen darauf hin, dass die Items nicht ausschließlich einer einzelnen latenten Dimension zugeordnet sind, sondern vielmehr Aspekte sowohl von *Verbundenheit* als auch von *Lernen* erfassen. Solche Muster können die Validität schwächen und die konzeptionelle Klarheit der zugrunde liegenden Konstrukte beeinträchtigen [42]. So können Items mit hohen Ladungen auf mehr als einen Faktor zusätzliche Messfehler verursachen und die Interpretation der Skalen erschweren [43]. Werden Kreuzladungen nicht berücksichtigt, kann die resultierende Faktorenstruktur an Stabilität verlieren, wodurch das Risiko einer Fehlspezifikation des Modells

steigt und möglicherweise ungenaue oder fehlerhafte theoretische Interpretationen entstehen [44].

Diese Überlegungen legen nahe, dass eine weitere Präzisierung des Messmodells erforderlich ist. Ein Ansatz zur Weiterentwicklung des Instruments besteht in der Entwicklung einer Kurzform der Skala.

Im Einklang mit etablierten Prinzipien der Skalenkonstruktion wird grundsätzlich eine sparsame und klar strukturierte Faktorenlösung bevorzugt [32], [45]. Dementsprechend sollte bei der Auswahl der Items der Fokus auf Items liegen, die eindeutig auf einen einzelnen vorgesehenen Faktor laden, sodass jedes Item seinen zugewiesenen Inhaltsbereich möglichst präzise abbildet [45]. Entsprechend dienen explorative Faktorenanalysen in der Regel als Grundlage für die Entfernung von Items mit schwachen Hauptladungen oder ausgeprägten Ladungen auf mehreren Faktoren. Items, die wiederholt Kreuzladungen aufweisen, werden daher häufig als ungeeignet für die Beibehaltung in der Skala angesehen [42], [43].

Der Empfehlung von Barnard-Brak und Shiu folgend, die Faktorenstruktur mithilfe explorativer Faktorenanalysen und einer Analyse auf Itemebene anzupassen, um das Instrument weiterzuentwickeln, wendeten Cho und Demmans Epp diesen Ansatz auf Daten aus Online-Lernumgebungen an [26], [41]. Daraus entstand eine englischsprachige Kurzform der CCS (CCS-SF), die in der Literatur weiterhin validiert wurde [41], [46]. Diese Version wurde jedoch auf Grundlage anderer Datensätze und in einem abweichenden Kontext entwickelt, nämlich in Online-Lernumgebungen.

Da die Validierung der CCS-D bereits Unterschiede in den Itemladungen im Vergleich zur ursprünglichen CCS gezeigt hatte [4], erschien es notwendig, eine Kurzform auf Grundlage des vorliegenden Datensatzes zu entwickeln.

Da Übersetzungsprozesse subtile semantische Veränderungen hervorrufen können, die sich auf die Interpretation der Items auswirken, ist eine empirische Validierung im deutschen Kontext von besonderer Bedeutung.

Die Verkürzung einer Skala bringt jedoch methodische Herausforderungen mit sich. Es besteht kein methodischer Konsens darüber, wie viele Items bei der Entwicklung einer Kurzform optimalerweise beibehalten werden sollten [47]. Eine zu starke Reduktion der Items kann die Inhaltsvalidität und Reliabilität beeinträchtigen, wenn die verschiedenen Bereiche des Konstrukts nicht ausreichend abgebildet werden [48]. Daher sollte die Auswahl der Items nach sorgfältigen und transparenten Verfahren erfolgen, um mögliche Einbußen bei der Qualität von Erhebungen zu minimieren [37].

2. Methoden

2.1. Protokoll der systematischen Itemreduktion des CCS-D-Fragebogens

Ausgehend von den oben dargestellten Begründungen und Herausforderungen wurde auf Grundlage der bereits übersetzten Version des Instruments ein systematisches Verfahren zur Itemreduktion entwickelt und umgesetzt. Ziel war es, den Fragebogen auf insgesamt zehn Items zu verkürzen, fünf pro Subskala, ohne dabei die psychometrische Qualität, theoretische Grundlage oder die inhaltliche Abdeckung der Subskalen wesentlich zu beeinträchtigen.

Die Entscheidung, eine Kurzversion mit zehn Items zu entwickeln, basierte auf mehreren methodischen und konzeptionellen Überlegungen.

Methodische Forschung deutet darauf hin, dass weniger als vier Items pro Faktor zu einer unzureichenden internen Reliabilität führen können [47], [49]. Darüber hinaus betont die Literatur zur Skalenentwicklung, dass Instrumente ausreichend viele Items enthalten sollten, um den jeweiligen Inhaltsbereich des Konstrukts angemessen abzubilden und eine Unterrepräsentation des Konstrukts zu vermeiden [50]. Folglich betonen Empfehlungen zur Skalenreduktion die Bedeutung eines ausgewogenen Verhältnisses zwischen Kürze, einer angemessenen Abdeckung des Konstrukts und der Reliabilität [33], [35], [37], [51].

Unter den Annahmen der klassischen Testtheorie steigt die interne Konsistenz tendenziell mit der Anzahl der Items, insbesondere bei der Verwendung von Cronbachs Alpha, sofern die Items dasselbe zugrunde liegende Konstrukt messen [52], [53], [54]. Die Beibehaltung einer etwas größeren Anzahl von Items in einer Zehn-Item-Version im Vergleich zu einer stärker verkürzten Acht-Item-Version trägt daher zu stabileren Messeigenschaften bei.

Darüber hinaus ermöglicht eine Skala dieser Länge eine umfassendere inhaltliche Abdeckung des zugrunde liegenden Konstrukts. Dies ist für die Konstruktvalidität von wesentlicher Bedeutung, da eine Reduktion der Items zu einer unzureichenden Abbildung relevanter Aspekte führen kann [48], [50]. Insbesondere bei multidimensionalen Konstrukten wie der Klassengemeinschaft kann eine zu starke Reduktion der Items die konzeptionelle Breite des Instruments einschränken und seine Fähigkeit schwächen, das zugrunde liegende Konstrukt angemessen abzubilden [48], [50].

Weiterhin stellt die Zehn-Item-Version einen praktischen Kompromiss zwischen Messgenauigkeit und Effizienz dar, da sie die Bearbeitungszeit der Teilnehmenden deutlich reduziert und gleichzeitig die psychometrische Qualität des Instruments bewahrt [51].

Der systematische Reduktionsprozess wurde in einem dreistufigen Verfahren durchgeführt. Das Vorgehen orientierte sich dabei sowohl an der iterativen Methodik von

Cho und Demmans Epp bei der Validierung ihrer CCS-Kurzform [41] als auch an den Leitlinien von Hinkin [45]. Das zur Reduktion zu verwendende Protokoll wurde zu Beginn des Überarbeitungsverfahrens festgelegt, um einen durch die Forschenden bedingten Selektionsbias zu minimieren. Als Ausgangspunkt diente der vollständige CCS-D-Datensatz von Knof und Shiozawa [4]. Zur Untersuchung der Faktorenstruktur wurde eine Hauptkomponentenanalyse (PCA) mit Varimax-Rotation durchgeführt, während die interne Konsistenz der Skalen mithilfe von Cronbachs Alpha bestimmt wurde. Anschließend wurde ein iterativer Reduktionsprozess gemäß dem vorab definierten Protokoll eingeleitet. Ziel des Verfahrens war die Optimierung des Konstrukts unter gleichzeitiger Wahrung der inhaltlichen Abdeckung und nicht die Maximierung der Koeffizienten der internen Konsistenz.

In jeder Iteration begann das Verfahren mit Schritt 1: Alle Items, die Kreuzladungen aufwiesen, also sekundäre Ladungen ≥.30, oder deren Hauptladungen auf einem nicht vorgesehenen Faktor lagen, wurden gleichzeitig entfernt. Hauptladungen auf einem nicht vorgesehenen Faktor wurden als Items definiert, die stärker auf einen anderen Faktor luden als auf die theoretisch erwartete Dimension, basierend auf der ursprünglichen Skalenstruktur und dem oben beschriebenen theoretischen Rahmen von Rovai. Diese Items bilden kein einzelnes latentes Konstrukt eindeutig ab und können daher die Validität der Skala beeinträchtigen [43], [53]. Wurden in Schritt 1 ein oder mehrere Items entfernt, wurde unmittelbar eine neue PCA durchgeführt und Cronbachs Alpha neu berechnet, wodurch eine neue Iteration begann. Erfüllte kein Item in Schritt 1 die Kriterien für eine Entfernung, wurde Schritt 2 angewendet. In diesem Schritt wurde dasjenige Item entfernt, dessen Ausschluss zu einer Verbesserung der internen Konsistenz führte, gemessen anhand von Cronbachs Alpha bei Entfernung des jeweiligen Items. Nach jeder Entfernung in Schritt 2 wurden erneut eine PCA durchgeführt und Cronbachs Alpha berechnet, bevor das Verfahren in der nächsten Iteration zu Schritt 1 zurückkehrte. Führten weder Schritt 1 noch Schritt 2 zur Entfernung eines Items, wurde Schritt 3 durchgeführt. Dabei wurden gleichzeitig alle Items mit den niedrigsten Ladungen auf ihren theoretisch zugeordneten Faktoren entfernt, um die übergeordnete Konstruktvalidität der Skala zu bewahren und eine angemessene Repräsentation der zugrunde liegenden theoretischen Dimensionen aufrechtzuerhalten [43], [53]. Wie in den vorherigen Schritten wurden nach jeder Entfernung eine PCA und eine Neuberechnung von Cronbachs Alpha durchgeführt, bevor der Prozess erneut bei Schritt 1 begann. Dieses iterative Verfahren wurde fortgesetzt, bis keine weiteren Items gemäß den vorab definierten Kriterien entfernt werden konnten. Dadurch entstand eine psychometrisch optimierte Kurzform mit einer gleichmäßigen Repräsentation beider Subskalen mit jeweils fünf Items pro Faktor, wodurch die strukturelle Symmetrie des Messmodells der Klassengemeinschaft gewährleistet ist. Die Validierung der Kurzform wurde unter Verwendung einer unabhängigen Stichprobe gemäß den Empfehlungen guter wissenschaftlicher Praxis durchgeführt [33], [39], [48], [55]. Dieser Schritt ist wichtig, da die Antworten auf Items durch den umgebenden Kontext der Befragung beeinflusst werden können [33]. Daher könnte die Validierung des verkürzten Instruments anhand desselben Datensatzes, der für die Itemauswahl verwendet wurde, zu Verzerrungen führen.

2.2. Datenerhebung

Am Ende des Sommersemesters 2025 wurden Medizinstudierende des ersten Semesters an der Medizinischen Fakultät der Universität Tübingen mithilfe eines Fragebogens befragt. Zu diesem Zeitpunkt hatten die Teilnehmenden alle Lehrveranstaltungen ihres ersten Semesters abgeschlossen. Die Rekrutierung erfolgte persönlich durch den Erstautor und die Zweitautorin. Vor der Teilnahme willigten alle Studierenden nach entsprechender Aufklärung schriftlich in die Teilnahme ein. Sie wurden über den Zweck der Studie, mögliche Vorteile und Risiken, Rechte und Pflichten sowie die Freiwilligkeit der Teilnahme informiert. Die Daten wurden pseudonymisiert erhoben, und alle Teilnehmenden stimmten der Veröffentlichung anonymisierter Daten zu. Darüber hinaus wurden sie informiert, dass die Einwilligung jederzeit ohne negative Konsequenzen widerrufen werden konnte. Ein Ethikvotum für die Studie wurde von der Ethikkommission der Eberhard Karls Universität Tübingen erteilt (Nr.: 086/2022B02).

2.3. Datenanalyse

Analog zur Validierung der CCS-D wurde dasselbe Validierungsverfahren angewendet [4]. Es wurde eine deskriptive Analyse durchgeführt, um die Stichprobe darzustellen und die Skalen- sowie Itemmerkmale des CCS-D-SF zu bestimmen. Cronbachs Alpha verwendet, um die interne Konsistenz der Subskalen sowie die Konsistenz des gesamten Fragebogens zu bewerten. Werte von .70 oder höher gelten als zufriedenstellend [54].

Die Konstruktvalidität der Zwei-Faktoren-Struktur der CCS-D-SF wurde mithilfe einer konfirmatorischen Faktorenanalyse (CFA) unter Verwendung der Maximum-Likelihood-Methode (ML) untersucht.

In der CFA wurden die beobachteten Variablen, bestehend aus fünf Items zur *Verbundenheit* und fünf Items zum *Lernen*, als Indikatoren ihrer jeweiligen latenten Variablen *Verbundenheit* und *Lernen* spezifiziert. Die Modellgüte wurde anhand internationaler Richtlinien bewertet [56], [57], [58].

Folgende Indizes wurden verwendet:

- Normierter Chi-Quadrat-Güteindikator (χ^2/df): eine akzeptable Modellgüte wird durch Werte <3 [59].
- Comparative Fit Index (CFI) und Tucker–Lewis Index (TLI): gute Modellpassung bei ≥.95 [60].
- Root Mean Square Error of Approximation (RMSEA): Werte zwischen .05 und .08 gelten als akzeptabel [61], [62], [63].

- Standardized Root Mean Square Residual (SRMR): akzeptable Modellgüte bei Werten <.08 [60].

Zur weiteren Untersuchung der Robustheit des Modells und zur Exploration zugrunde liegender datenbasierter Muster, die durch die CFA möglicherweise nicht erfasst wurden, wurde zusätzlich eine PCA durchgeführt. Die PCA erfolgte mit Varimax-Rotation und Kaiser-Normalisierung. Items mit Faktorladungen von ≥|.30| auf eine oder beiden Dimensionen wurden beibehalten [59], [67]. Dieser ergänzende Ansatz folgte Empfehlungen, die die kombinierte Anwendung von CFA und explorativen Verfahren zur Sicherstellung der Modellrobustheit betonen [64], [65], [66]. Die Datenaufbereitung und die deskriptiven Analysen wurden mit IBM SPSS Statistics 30 (IBM Corp., Armonk, NY) durchgeführt. Die CFA wurde mit IBM SPSS AMOS 26 (IBM Corp., Armonk, NY) durchgeführt.

3. Ergebnisse

3.1. Stichprobe

Insgesamt wurden *N*=178 Medizinstudierende des ersten Semesters an der Medizinischen Fakultät Tübingen befragt. Maximal hätten 205 Studierende in die Untersuchung eingeschlossen werden können, was einer Rücklaufquote von 86,8% entspricht. Um die Bedingungen für die Nutzung von AMOS in der weiteren Analyse zu erfüllen, wurden alle Teilnehmer mit fehlenden Daten zu den CCS-D-SF-Items ausgeschlossen. Dadurch verblieben 171 Befragte.
Die Studienteilnehmenden waren zu 68,4% weiblich, was der Geschlechterverteilung der zugelassenen Studierenden entsprach. Das Alter reichte von 18 bis 29 Jahren, das Durchschnittsalter betrug 20,43 Jahre (*SD*=1,97). Die Mehrheit der Teilnehmenden waren Medizinstudierende (87,1%), gefolgt von Zahnmedizinstudierenden (12,9%).

3.2. Systematische Itemreduktion des CCS-D-Fragebogens

Die Validierung der schrittweisen Itemreduktion für die Kurzform basierte auf der PCA des CCS-D-Datensatzes, die eine gute interne Konsistenz zeigte (Cronbach's α=.87; siehe Anhang 1, Tabelle S2) [4].
Auf Grundlage dieser Analyse wurde Schritt 1 des Itemreduktionsprotokolls durchgeführt. Alle Items mit ausgeprägten Kreuzladungen (Items 9, 10, 14 und 16) wurden entfernt. Anschließend wurde eine PCA durchgeführt und Cronbachs Alpha erneut berechnet, um die Grundlage für die nächste Iteration zu schaffen. Dies führte zur ersten Iteration der CCS-D-SF mit einem Cronbachs Alpha von .84 (siehe Anhang 1, Tabelle S3).
In der ersten Iteration der CCS-D-SF zeigten keine Items Kreuzladungen oder Hauptladungen auf einen nicht vorgesehenen Faktor. Entsprechend wurde Schritt zwei ausgeführt, der zur Entfernung des Items 6 führte, da dieser Ausschluss zur Erhöhung von Cronbach's Alpha führte. Die daraufhin durchgeführten PCA und Reliabilitätsanalysen führten zur CCS-D-SF Iteration 2 (Cronbach's Alpha=.84) (siehe Anhang 1, Tabelle S4).
In CCS-D-SF Iteration 2 wurden die Items 5 und Item 17 aufgrund ihrer Kreuzladungen entfernt. Die folgende PCA und Berechnung von Cronbach's Alpha führte zu CCS-D-SF Iteration 3 (Cronbach's Alpha=.82) (siehe Anhang 1, Tabelle S5).
In Iteration 3 wurde Item 1 entfernt, da es Ladungen auf mehr als einen Faktor aufwies. Die folgende PCA und Berechnung von Cronbach's Alpha führte zu CCS-D-SF Iteration 4 (Cronbach's Alpha=.80) (siehe Anhang 1, Tabelle S6).
In Iteration 4 wiesen keine Items Kreuzladungen oder Hauptladungen auf einen nicht vorgesehenen Faktor auf. Schritt 2 zeigte keine Verbesserungen von Cronbach's Alpha, entsprechend wurde Schritt 3 ausgeführt, was zur Entfernung der Items 7 und 2 führte, da diese die niedrigsten Faktorladungen aufwiesen.
Die resultierende 10-Item Version enthielt fünf Items pro Subskala. Eine abschließende PCA und Berechnung von Cronbach's Alpha führte zu CCS-D-SF Iteration 5 (Cronbach's Alpha=.79) (siehe Anhang 1, Tabelle S7).
Um die Nutzbarkeit zu verbessern, wurden die Items in abwechselnder Reihenfolge ihrer Subskalen neu geordnet und entsprechend nummeriert, dem originalen CCS-Format folgend (siehe Tabelle 1). Die finale CCS-D-SF enthält zwei Subskalen: *Verbundenheit* (fünf ungerade-nummerierte Items) und *Lernen* (fünf gerade-nummerierte Items). Jedes Item wird auf einer fünfstufigen Likert-Skala bewertet: Wie im originalen CCS Instrument gilt für die Items 1, 3, 5, 7, 9, folgende Skala: *stimme vollständig zu=4, stimme zu=3, neutral=2, stimme nicht zu=1, stimme überhaupt nicht zu=0;* für die Items 2, 4, 6, 8, 10 lautete die Bewertungsskala: *stimme vollständig zu=0, stimme zu=1, neutral=2, stimme nicht zu=3, stimme überhaupt nicht zu=4* [3]. Für die weitere Analyse müssen diese Item-Scores invertiert werden. Diese Items sind in den Tabellen und Abbildungen dieser Arbeit mit einem „i" gekennzeichnet.

3.3. Validierung der CCS-D-SF

3.3.1. Skalenmerkmale und interne Konsistenz

Die deskriptiven Statistiken für die einzelnen Items (siehe Tabelle 2) sowie für die gesamte CCS-D-SF und jede Subskala sind unten dargestellt (siehe Tabelle 3). Die gesamte CCS-D-SF (Cronbach's Alpha=.81) sowie die Subskala *Verbundenheit* (Alpha=.86) zeigten eine gute interne Konsistenz. Die Subskala *Lernen* weist einen Cronbach's Alpha von .65 auf.

3.3.2. Konfirmatorische Faktorenanalyse

Das CFA-Modell für das 2-Faktoren-Modell ist in Abbildung 1 dargestellt.

Tabelle 1: Finale Version der CCS-D-SF

1	Ich fühle mich mit anderen Studierenden dieses Kurses verbunden.
2	Ich habe das Gefühl, dass es schwierig ist Hilfe zu erhalten, wenn ich eine Frage habe.
3	Ich vertraue den anderen Kursteilnehmern.
4	Ich fühle mich unsicher meine Wissenslücken zu offenbaren.
5	Ich habe das Gefühl, dass ich mich auf die anderen Kursteilnehmer verlassen kann.
6	Ich denke der Kurs bringt nur einen mittelmäßigen Lernerfolg.
7	Ich habe das Gefühl, dass andere Kursteilnehmer sich auf mich verlassen.
8	Ich habe das Gefühl, dass meine Lernbedürfnisse nicht erfüllt werden.
9	Ich bin zuversichtlich, dass die anderen Kursteilnehmer mich unterstützen.
10	Ich habe das Gefühl, dass dieser Kurs die Lernmotivation nicht fördert.

Tabelle 2: Deskriptive Statistik der CCS-D-SF-Items

Deskriptive Statistik					
	N	Minimum	Maximum	Mittelwert	Standardabweichung
Item 1	171	0	4	3.15	.994
Item 2 i	171	0	4	2.84	1.002
Item 3	171	0	4	3.09	.900
Item 4 i	171	0	4	2.39	1.113
Item 5	171	0	4	3.09	.867
Item 6 i	171	0	4	2.64	.956
Item 7	171	0	4	2.70	.874
Item 8 i	171	0	4	2.92	.985
Item 9	171	0	4	3.04	.900
Item 10 i	171	0	4	2.81	.894
Valid N (listwise)	171				

Tabelle 3: Deskriptive Statistik der CCS-D-SF-Subskalen

Deskriptive Statistik					
	N	Minimum	Maximum	Mittelwert	Standardabweichung
CCS-D-SF Gesamt	171	5.00	39.00	28.67	5.787
CCS-D-SF Verbundenheit	171	2.00	20.00	15.06	3.630
CCS-D-SF Lernen	171	3.00	20.00	13.60	3.269
Valid N (listwise)	171				

Das CFA-Modell zeigte übergreifend eine akzeptable Modellanpassung. Die zwei latenten Faktoren waren moderat korreliert (r=.51). Das normierte Chi-Quadrat (χ^2 [34, n=171]=73.529, p<.001; χ^2/df=2.163) zweigt eine ausreichende Passung. Die RMSEA (.083) überschritt den Grenzwert für einen akzeptablen Modellfit, während der SRMR (.059) auf einen guten Modellfit hinwies. Die vergleichenden Fit-Indizes CFI (.93) und TLI (.91) lagen im akzeptablen Bereich, erreichten jedoch nicht den Schwellenwert für einen guten Modellfit.

3.3.3. Hauptkomponentenanalyse

Das Kaiser-Meyer-Olkin-Maß betrug .811 und wies damit auf eine relativ gute Eignung der Daten für eine Faktorenanalyse hin. Der Bartlett-Test auf Sphärizität war signifikant (*approx.* χ^2=581.215; *p*<.001), , was darauf hindeutet, dass die Daten keine Einheitsmatrix bilden und für eine Faktorenanalyse geeignet sind. Es wurden nur Faktoren mit Eigenwerten ≥1 beibehalten [68], [69]. Basierend auf dem Kaiser-Kriterium und dem Scree-Plot wurden zwei Faktoren mit Eigenwerten größer als 1 bebehalten, die zusammen 54,9% der Gesamtvarianz erklärten. Im mittels Varimax-Rotation rotierten Zwei-Faktoren-Modell (siehe Tabelle 4) luden die meisten Items primär auf einen einzelnen Faktor, mit Ausnahme von Item 10 („Ich habe das Gefühl, dass dieser Kurs die Lernmotivation nicht fördert."). Die zugrunde liegende Faktorenstruktur der CCS-D-SF kann somit bestätigt werden, wobei die ungerade-nummerierten Items hoch auf Faktor 1 und die gerade-nummerierten Items auf Faktor 2 luden. Diese repräsentieren die Dimensionen *Verbundenheit* und *Lernen*.

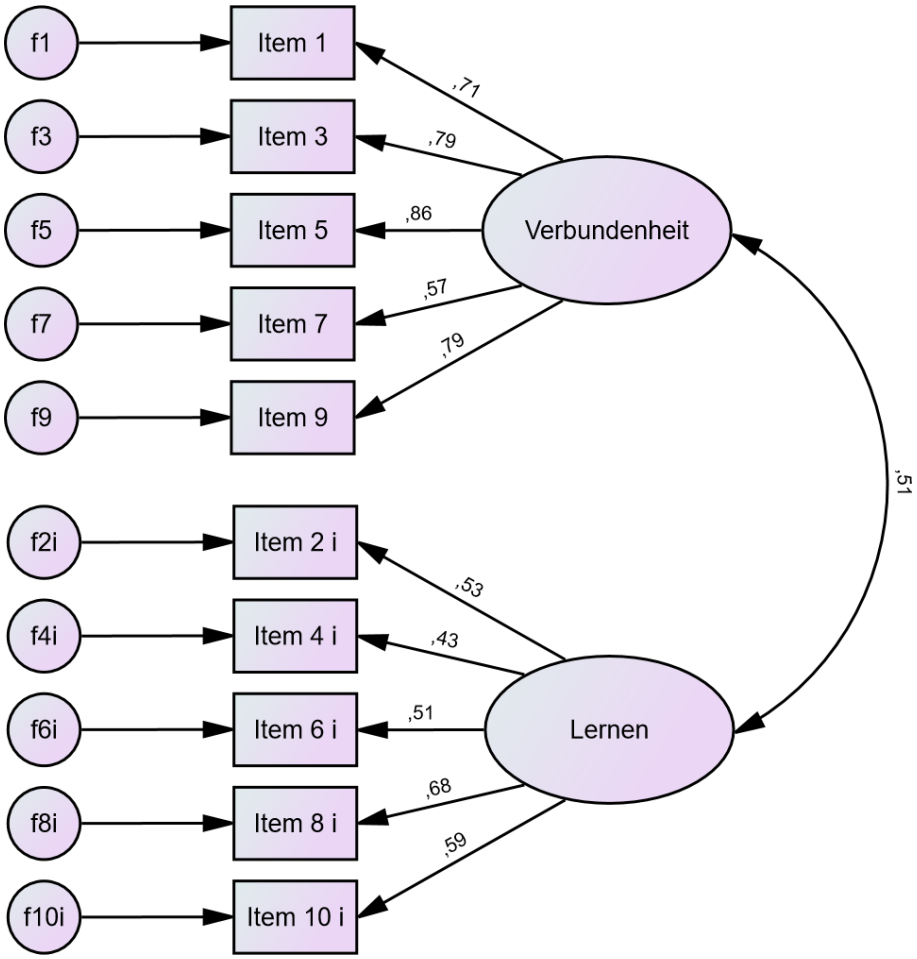


Abbildung 1: CCS-D-SF CFA-Modell

Tabelle 4: Hauptkomponentenanalyse

Rotierte Komponentenmatrix[a]		
	Komponenten	
	1	2
Item 1	.783	
Item 2 i		.565
Item 3	.816	
Item 4 i		.645
Item 5	.862	
Item 6 i		.661
Item 7	.637	
Item 8 i		.775
Item 9	.812	
Item 10 i	.343	.545

Extraktionsmethode: Hauptkomponentenanalyse.
Rotationsmethode: Varimax mit Kaiser-Normalisierung
a. Die Rotation ist in 3 Iterationen konvergiert.

4. Diskussion

Die CCS-D-SF weist eine gute Reliabilität auf und repliziert in der CFA die Faktorenstruktur der Originalversion. In einer abschließenden PCA konnte zudem sichergestellt werden, dass sich die Faktorenlösung nach der Entfernung der Items nicht verändert hatte [51].

Bei Betrachtung der Modellfit-Indizes zeigt die CCS-D-SF eine bessere Passung als die CCS-D.

Zwischen den Subskalen *Verbundenheit* und *Lernen* wurde eine moderate Korrelation festgestellt. Dies deutet darauf hin, dass sie einen erheblichen Anteil gemeinsamer Varianz aufweisen und gleichzeitig unterschiedliche Konstrukte repräsentieren. Die Beziehung weist jedoch darauf hin, dass die beiden Dimensionen der Klassengemeinschaft nicht vollständig unabhängig voneinander sind. Die geringfügig erhöhte RMSEA könnte teilweise auf die erhebliche Überschneidung zwischen *Verbundenheit* und *Lernen* zurückzuführen sein. Eine ähnlich hohe Korrelation zwischen den Subskalen wurde bereits in der Literatur berichtet [30].

Diese Beziehung steht im Einklang mit dem *Community-of-Inquiry*-Modell, das davon ausgeht, dass soziale Präsenz *(Verbundenheit)* und kognitive Präsenz (*Lern*engagement) dynamisch miteinander interagieren und dadurch Lernerlebnisse fördern [12], [70]. Studierende, die sich sozial in einen Kurs integriert fühlen, beteiligen sich mit höherer Wahrscheinlichkeit auch kognitiv am Lernprozess. Umgekehrt kann eine intensive kognitive Auseinandersetzung das Gefühl der Zugehörigkeit und der sozialen Verbundenheit verstärken [11]. Die Subskala *Verbundenheit* zeigte eine sehr gute interne Konsistenz, während *Lernen* lediglich eine knapp akzeptable Reliabilität aufwies. Die geringere Reliabilität der Subskala *Lernen* steht im Einklang mit früheren Studien, die die CCS untersuchten und eine niedrigere interne Konsistenz für lernbezogene Items berichteten [26], [27], [71]. Die reduzierte Anzahl an Items in der CCS-D-SF erfasst diese konzeptionelle Breite möglicherweise nicht vollständig, wodurch die interne Konsistenz geringer ausfallen und möglicherweise auch die Korrelation zwischen den Subskalen beeinflusst werden könnte. Darüber hinaus könnte *Lernen* stärker kontextabhängig sein als *Verbundenheit*, da die Wahrnehmung des Lernens durch die Art der Instruktionen, den Lehrstil und die Prüfungen beeinflusst wird.

Die Kreuzladung von Item 10 („Ich habe das Gefühl, dass dieser Kurs die Lernmotivation nicht fördert“) auf die Subskala *Verbundenheit* unterstützt diese Interpretation zusätzlich. Dieses Muster könnte die Art und Weise widerspiegeln, wie Studierende den Begriff „Kurs“ wahrnehmen, der sowohl als Inhalt und Struktur der Lehrveranstaltung, als soziale Gruppe der Teilnehmenden oder als Kombination aus beiden verstanden werden kann. Darüber hinaus können externe Faktoren wie die Persönlichkeit und das Auftreten der Lehrperson die Wahrnehmung der Studierenden hinsichtlich sozialer Präsenz und Klassengemeinschaft beeinflussen und dadurch sowohl *Verbundenheit* als auch *Lernen* beeinflussen [72], [73]. Auch die Erwartungen der Studierenden an zukünftige Interaktionen mit ihren Mitstudierenden in weiteren Lehrveranstaltungen könnten ihre Motivation zur Entwicklung eines Gemeinschaftsgefühls beeinflussen und somit zur beobachteten Korrelation zwischen den Subskalen beitragen [74].

Aus pädagogischer Perspektive hat diese wechselseitige Abhängigkeit zwischen *Verbundenheit* und *Lernen* wichtige Implikationen. Lehrstrategien, die gezielt den sozialen Zusammenhalt fördern, beispielsweise Gruppenprojekte, Peer-Mentoring oder strukturierte Reflexionen, können nicht nur die *Verbundenheit* stärken, sondern auch die Lernergebnisse verbessern [12], [24]. Ebenso könnten Curriculumdesigns, die die kollaborative Konstruktion von Wissen in den Mittelpunkt stellen, sowohl das soziale Zugehörigkeitsgefühl als auch das kognitive Engagement stärken. Die CCS-D-SF könnte daher als diagnostisches Instrument für Lehrende dienen, um Ungleichgewichte zwischen diesen beiden Dimensionen zu identifizieren und gezielte Interventionen zur Verbesserung der Klassengemeinschaft anzuleiten.

Es gibt mehrere Limitationen dieser Studie, die berücksichtigt werden sollten. Bei Studierendengemeinschaften kann eine Homogenisierung beobachtet werden, die durch ähnliche soziale und persönliche Herausforderungen, einschließlich der mit dem Übergang in die Hochschulbildung verbundenen Unsicherheiten, begünstigt wird [75]. Studierende im ersten Semester konzentrieren sich typischerweise auf den Aufbau persönlicher Beziehungen und die Etablierung eines für ihr Studium relevanten beruflichen Netzwerks [76]. Dies kann beeinflussen, wie Studierende *Verbundenheit* und *Lernen wahrnehmen* oder priorisieren, da soziale Integration und der Aufbau von Netzwerken vorübergehend eine größere Bedeutung haben können als akademisches Engagement oder lernorientierte Interaktionen. Im Gegensatz zur vorliegenden Studie, in der das Gemeinschaftsgefühl über einen gesamten Studiengang hinweg gemessen wurde, konzentrierte sich der Großteil der bisherigen Forschung unter Verwendung der CCS auf einzelne Kurse oder Unterrichtseinheiten [77].

Bei der Interpretation der Ergebnisse des Skalenreduktionsprozesses sollten mehrere methodische Aspekte berücksichtigt werden. Faktorenanalytische Verfahren werden häufig zur Verkürzung von Skalen eingesetzt, da Faktorladungen die Stärke der Beziehung zwischen Items und den zugrunde liegenden latenten Dimensionen anzeigen [33], [78]. Dementsprechend stützen sich viele Reduktionsverfahren auf psychometrische Kennwerte wie Faktorladungen, Item-Gesamtwert-Korrelationen und Reliabilitätsschätzungen [35]. In der vorliegenden Studie basierte die Weiterentwicklung der Skala jedoch nicht ausschließlich auf Faktorladungen. Als ergänzendes Kriterium wurde die interne Konsistenz des Gesamtinstruments berücksichtigt, um die Stabilität des Messmodells während des gesamten Reduktionsprozesses sicherzustellen [50], [52]. Eine ausschließliche Orientierung an statistischen Kriterien kann jedoch zu verkürzten Skalen führen, die zwar psychometrisch homogen, konzeptionell jedoch eng gefasst sind und deren inhaltliche Abdeckung

zu restriktiv ist [34], [50]. Zudem besteht die Möglichkeit, dass solche Instrumente geringfügig andere Aspekte des Konstrukts erfassen als die Langform [33]. Um diesen Bedenken zu begegnen, wurde darauf geachtet, eine gleichmäßige Repräsentation der beiden zugrunde liegenden Dimensionen der Klassengemeinschaft, *Verbundenheit* und *Lernen*, zu erhalten [3].

Die Autoren erkennen an, dass der vorliegende Ansatz eine von mehreren derzeit diskutierten methodischen Strategien zur Skalenreduktion darstellt. Kurzformen von Instrumenten werden häufig entwickelt, indem Items bestehender längerer Fragebögen auf Grundlage psychometrischer Kennwerte schrittweise entfernt werden [32]. Obwohl diese Ansätze vergleichsweise einfach umzusetzen sind, berücksichtigen sie häufig nur eine begrenzte Anzahl psychometrischer Kriterien und können daher zu Schwächen wie unzureichenden Faktorenstrukturen, mangelnder Reliabilität oder fehlender Messinvarianz führen [79]. In letzter Zeit wurden algorithmische Ansätze vorgeschlagen, die bei der Entwicklung von Kurzformen vielversprechende Ergebnisse gezeigt haben [79], [80]. Heuristische Optimierungsverfahren können jedoch nicht garantieren, dass eine einzelne optimale Lösung identifiziert wird, und können bei wiederholten Analysen zu unterschiedlichen Kombinationen von Items führen [32]. Diese Entwicklungen verdeutlichen, dass die Verkürzung von Skalen nicht als rein mechanisches Optimierungsverfahren verstanden werden sollte, sondern vielmehr als kontextabhängiger Prozess, der sowohl statistische als auch theoretische Überlegungen erfordert [79], [80].

Da datenbasierte Verfahren zur Itemreduktion gegenüber stichprobenspezifischen Merkmalen empfindlich sein können, wurde die resultierende Kurzform mithilfe einer CFA in einer unabhängigen Validierungsstichprobe untersucht [45], [55]. Dieser Ansatz ermöglicht zudem die Beurteilung der Stabilität und Replizierbarkeit der Faktorenstruktur und berücksichtigt, dass Evidenz der Validität, die für Langformen von Skalen gewonnen wurde, nicht zwangsläufig auf verkürzte Versionen übertragbar ist [45], [55]. Alternative Strategien, wie beispielsweise eine Split-Sample-Validierung innerhalb des ursprünglichen Datensatzes, hätten das Risiko eines stichprobenspezifischen Overfittings verringern können. Die Umsetzung eines Split-Sample-Ansatzes hätte jedoch die für die Itemauswahl und Faktorenanalyse verfügbare effektive Stichprobengröße reduziert. Dies ist insbesondere vor dem Hintergrund problematisch, dass Faktorenanalysen bei kleineren Stichproben instabil werden können [63], [81]. Die vorliegende Studie umfasste keine externen Validierungskriterien, die einen direkten Vergleich der Validität der Kurzform ermöglicht hätten. Ohne solche externen Bezugskriterien kann letztlich nicht abschließend beurteilt werden, ob dieses Instrument tatsächlich dasselbe zugrunde liegende Konstrukt erfasst [33]. Verbesserungen in der Modellgüte nach der Skalenreduktion sollten nicht grundsätzlich als Hinweis auf eine überlegene Messqualität interpretiert werden, da Fit-Indizes empfindlich auf die Modellkomplexität, die Länge des Instruments und die Homogenität der Items reagieren [82], [83].

Zukünftige Forschung sollte daher die Zusammenhänge zwischen der CCS-D-SF und theoretisch verwandten Konstrukten untersuchen, um die Konstruktvalidität der Skala weiter zu belegen.

5. Schlussfolgerung

Diese Studie zeigt die systematische Entwicklung und Validierung einer Kurzform der deutschen Version der Classroom Community Scale (CCS-D-SF).

Die CCS-D-SF zeigte eine gute Reliabilität, die mit bisherigen Befunden aus der Literatur vergleichbar ist. Die konfirmatorische Faktorenanalyse bestätigte die Zwei-Faktoren-Struktur der Skala, wobei die Modellfit-Indizes sogar etwas besser ausfielen als die für die Langversion der CCS-D berichteten Werte. Die Faktorenstruktur wurde zudem durch die Hauptkomponentenanalyse gestützt.

Insgesamt stellt die CCS-D-SF ein praktisches und psychometrisch fundiertes Instrument zur Messung des Gemeinschaftsgefühls bei Medizinstudierenden im ersten Studienjahr dar. Durch ihre Kürze reduziert sie die Belastung der Teilnehmenden, während die Messqualität erhalten blieb oder sogar verbessert werden konnte.

Zukünftige Forschung sollte sich auf die Anwendung der CCS-D-SF in deutschsprachigen Ländern sowie auf die weitere Untersuchung ihrer Eignung in unterschiedlichen Bildungskontexten konzentrieren.

Tagungsvortrag

Auszüge aus diesem Manuskript wurden auf der Jahrestagung der DACH-Gesellschaft für Medizinische Ausbildung (GMA) 2026 in Dresden vorgestellt [84].

ORCIDs der Autor*innen

- Harald Knof: [0000-0002-0942-653X]
- Thomas Shiozawa: [0000-0002-7112-1016]

Danksagung

Die Autoren möchten Prof. Alfred P. Rovai von der Regent University (Virginia Beach, VA, USA) für die Genehmigung zur Nutzung und Übersetzung seines CCS-Instruments danken.

Des Weiteren danken die Autoren Prof. Dr. Peter Martus für die statistische Beratung.

Den Studierenden der Medizinischen Fakultät Tübingen sind die Autoren dankbar für ihre Teilnahme an dieser Studie.

Die Autoren danken für die Unterstützung durch den Open-Access-Publikationsfonds der Universität Tübingen.

Interessenkonflikt

Die Autoren erklären, dass keine Interessenkonflikte bestehen. Allein die Autoren sind für den Inhalt und das Schreiben dieses Artikels verantwortlich. Die Autoren geben an, dass die Forschung in Abwesenheit jeglicher kommerzieller oder finanzieller Beziehungen durchgeführt wurde, die als potenzieller Interessenkonflikt ausgelegt werden könnten.

Anhänge

Verfügbar unter https://doi.org/10.3205/zma001891

1. Anhang_1.pdf (157 KB)
   Zusätzliche Tabellen

Literatur

1. Rosen MA, DiazGranados D, Dietz AS, Benishek LE, Thompson D, Pronovost PJ, Weaver SJ. Teamwork in healthcare: Key discoveries enabling safer, high-quality care. Am Psychol. 2018;73(4):433-450. DOI: 10.1037/amp0000298
2. Van Leuven AJ, Szymik BG, Ramsey LM, Hesse DW. A multi-year evaluation of medical student performance on and perceptions of collaborative gross anatomy laboratory examinations. Anat Sci Educ. 2023;16(3):557-566. DOI: 10.1002/ase.2223
3. Rovai AP. Development of an instrument to measure classroom community. Internet High Educ. 2002;5(3):197-211. DOI: 10.1016/s1096-7516(02)00102-1
4. Knof H, Shiozawa T. Validation of the German form of Classroom-Community-Scale (CCS-D). GMS J Med Educ. 2025;42(1):Doc13. DOI: 10.3205/zma001737
5. Dillenbourg P. What do you mean by 'collaborative learning'? In: Dillenbourg P, editor. Collaborative Learning: Cognitive and Computational Approaches. Advances in learning and instruction series. 1st ed. London: Pergamon; 1999. (Advances in learning and instruction series). p.1-15.
6. Sevenhuysen S, Haines T, Kiegaldie D, Molloy E. Implementing collaborative and peer-assisted learning. Clin Teach. 2016;13(5):325-331. DOI: 10.1111/tct.12583
7. O'Brien BC, Forrest K, Wijnen☐Meijer M, ten Cate O. A Global View of Structures and Trends in Medical Education. In: Swanwick T, Forrest K, O'Brien BC, editors. Understanding Medical Education. 3rd ed. Hoboken: Wiley-Blackwill; 2018. p.7-22. DOI: 10.1002/9781119373780.ch2
8. Almajed A, Winning T, Skinner V, Peterson R. Medical and Dental Students' Perceptions and Experiences of Collaborative Learning: A Systematic Review. JBI Libr Syst Rev. 2011;9(32 Suppl):1-13. DOI: 10.11124/01938924-201109321-00001
9. Skinner VJ, Braunack-Mayer A, Winning TA. The Purpose and Value for Students of PBL Groups for Learning. Interdiscip J Problem-Based Learn. 2015;9(1). DOI: 10.7771/1541-5015.1499
10. Pluta WJ, Richards BF, Mutnick A. PBL and beyond: trends in collaborative learning. Teach Learn Med. 2013;25 Suppl 1:S9-S16. DOI: 10.1080/10401334.2013.842917
11. Chatterjee R, Correia A-P. Online Students' Attitudes Toward Collaborative Learning and Sense of Community. Am J Distance Educ. 2019;34(1):53-68. DOI: 10.1080/08923647.2020.1703479
12. Garrison DR, Anderson T, Archer W. The first decade of the community of inquiry framework: A retrospective. Internet High Educ. 2010;13(1-2):5-9. DOI: 10.1016/j.iheduc.2009.10.003
13. Halic O, Lee D, Paulus T, Spence M. To blog or not to blog: Student perceptions of blog effectiveness for learning in a college-level course. Internet High Educ. 2010;13(4):206-213. DOI: 10.1016/j.iheduc.2010.04.001
14. Lai KW. Knowledge construction in online learning communities: a case study of a doctoral course. Stud High Educ. 2013;40(4):561-579. DOI: 10.1080/03075079.2013.831402
15. Rovai AP. Building Sense of Community at a Distance. Int Rev Res Open Dis. 2002;3(1):1-16. DOI: 10.19173/irrodl.v3i1.79
16. Bruffee KA. Collaborative Learning: Higher Education, Interdependence, and the Authority of Knowledge. 1st ed. Baltimore: Johns Hopkins University Press; 1993.
17. Tinto V. Classrooms as Communities. J High Educ. 1997;68(6):599-623. DOI: 10.1080/00221546.1997.11779003
18. Sheets MF. Characteristics of Adult Education Students and Factors Which Determine Course Completion: A Review. New Horiz Adult Educ. 1992;6(1):3-18. DOI: 10.1002/nha3.10034
19. McMillan D, Chavis D. Sense of Community: A Definition and Theory. J Community Psychol. 1986;14(1):6-23. DOI: 10.1002/1520-6629(198601)14:1<6::AID-JCOP2290140103>3.0.CO;2-I
20. Hill JL. Psychological sense of community: Suggestions for future research. J Community Psychol. 1996;24(4):431-438. DOI: 10.1002/(Sici)1520-6629(199610)24:4<431::Aid-Jcop10>3.3.Co;2-Z
21. Rovai AP. Sense of community, perceived cognitive learning, and persistence in asynchronous learning networks. Internet High Educ. 2002;5(4):319-332. DOI: 10.1016/s1096-7516(02)00130-6
22. Shaffer C, Anundsen K. Creating community anywhere: Finding support and connection in a fragmented world. 1st ed. Los Angeles: Tarcher/Perigree; 1993.
23. Rheingold H. The virtual community: Finding commection in a computerized world. 1st ed. Boston: Addison-Wesley Longman Publishing Co. Inc.; 1993.
24. Strayhorn TL. College Students' Sense of Belonging - A Key to Educational Success for All Students. 2nd ed. New York: Routlege; 2018. DOI: 10.4324/9781315297293
25. Wang W, Gao Y, Pitts A, Dong L. A Bibliometric Analysis of Neighborhood Sense of Community. Sustainability. 2023;15(5):4183. DOI: 10.3390/su15054183
26. Barnard-Brak L, Shiu W. Classroom Community Scale in the blended learning environment: A psychometric review. Int J E-Learn. 2010;9(3):303-311.
27. Hur EH, Glassman M, Kim Y. Finding autonomy in activity: Development and validation of a democratic classroom survey. Educ Ass Eval Account. 2013;25(4):303-320. DOI: 10.1007/s11092-013-9173-y
28. Ahmady S, Kohan N, Bagherzadeh R, Rakshhani T, Shahabi M. Validity testing of classroom community scale in virtual environment learning: A cross sectional study. Ann Med Surg (Lond). 2018;36:256-260. DOI: 10.1016/j.amsu.2018.08.021
29. Yan J, Ding X, Xiong L, Liu E, Zhang Y, Luan Y, Qin L, Zhou C, Zhang W. Team-based learning: assessing the impact on anatomy teaching in People's Republic of China. Adv Med Educ Pract. 2018;9:589-594. DOI: 10.2147/AMEP.S169949
30. Abdeldayem M, Aldulaimi SH, Aldulaimi MLA. Virtual Learning and Students' Connectedness in the Time of Coronavirus. Int J Adv Sci. 2020;29(5):12634-12645.

31. Perrucci V, Cacciamani S, Khanlari A, Balboni G. The Classroom Community Scale in a Face-to-Face University Context: Factorial Structure and Convergent/Divergent Validity. Test Psychom Methodol Appl Psychol. 2022;29(3):297-307. DOI: 10.4473/TPM29.3.2

32. Franzen P, Arens AK, Greiff S, van der Westhuizen L, Fischbach A, Wollschläger R, Niepl C. Developing and Validating a Short-Form Questionnaire for the Assessment of Seven Facets of Conscientiousness in Large-Scale Assessments. J Pers Assess. 2022;104(6):759-773. DOI: 10.1080/00223891.2021.1998083

33. Stanton JM, Sinar EF, Balzer WK, Smith PC. Issues and Strategies for Reducing the Length of Self□Report Scales. Pers Psychol. 2002;55(1):167-194. DOI: 10.1111/j.1744-6570.2002.tb00108.x

34. Passarelli M, Casetta L, Rizzi L, Chiorri C, Cassina F, Voi S, Rocco D. Short and sweet: Comparing strategies for the reduction of questionnaires on self-criticism and social safeness while preserving construct validity. Int J Psychol. 2024;59(6):1234-1244. DOI: 10.1002/ijop.13249

35. Kilmen S, Bulut O. Shortening Psychological Scales: Semantic Similarity Matters. Educ Psychol Meas. 2025;85(5):910-934. DOI: 10.1177/00131644251319047

36. Yarkoni T. The Abbreviation of Personality, or how to Measure 200 Personality Scales with 200 Items. J Res Pers. 2010;44(2):180-198. DOI: 10.1016/j.jrp.2010.01.002

37. Alcaraz S, Viladrich C, Torregrosa M. Less time, better quality. Shortening questionnaires to assess team environment and goal orientation. Span J Psychol. 2013;16:E77. DOI: 10.1017/sjp.2013.74

38. Fraser BJ. Development of Short Forms of Several Classroom Environment Scales. J Educ Meas. 1982;19(3):221-227. DOI: 10.1111/j.1745-3984.1982.tb00130.x

39. Coste J, Guillemin F, Pouchot J, Fermanian J. Methodological approaches to shortening composite measurement scales. J Clin Epidemiol. 1997;50(3):247-252. DOI: 10.1016/s0895-4356(96)00363-0

40. Lewis K, McVay-Dyche J, Chen H, Soto T. Examining Sense of Community among Medical Professionals in an Online Graduate Program. J Educ Online. 2015;12(1). DOI: 10.9743/jeo.2015.1.8

41. Cho J, Demmans Epp C. Improving the Classroom Community Scale: Toward a Short-Form of the CCS. In: American Educational Research Association (AERA) Annual Meeting; 2019 Apr 5-9; Toronto, Canada. DOI: 10.3102/1432534

42. Brown TA. Confirmatory factor analysis for applied research. 2nd ed. New York (NY): The Guilford Press; 2015.

43. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis. 8th ed. Cheriton House: Cengage Learning; 2019.

44. Li Y, Wen Z, Hau KT, Yuan KH, Peng Y. Effects of Cross-loadings on Determining the Number of Factors to Retain. Struct Equation Model. 2020;27(6):841-863. DOI: 10.1080/10705511.2020.1745075

45. Hinkin TR. A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires. Organ Res Method. 1998;1(1):104-121. DOI: 10.1177/109442819800100106

46. Tackett M, Viel S, Manturuk K. A validation of the short-form classroom community scale for undergraduate mathematics and statistics students. J Uni Teach Learn Pract. 2023;20(1). DOI: 10.53761/1.20.01.08

47. Jokovic A, Locker D, Guyatt G. Short forms of the Child Perceptions Questionnaire for 11-14-year-old children (CPQ11-14): development and initial evaluation. Health Qual Life Outcomes. 2006;4:4. DOI: 10.1186/1477-7525-4-4

48. Smith GT, McCarthy DM, Anderson KG. On the sins of short-form development. Psychol Assess. 2000;12(1):102-111. DOI: 10.1037//1040-3590.12.1.102

49. Watson D, Clark LA. Measurement and mismeasurement of mood: recurrent and emergent issues. J Pers Assess. 1997;68(2):267-296. DOI: 10.1207/s15327752jpa6802_4

50. Clark LA, Watson D. Constructing validity: Basic issues in objective scale development. Psychol Assess. 1995;7(3):309-319. DOI: 10.1037/14805-012

51. Worthington RL, Whittaker TA. Scale Development Research: A Content Analysis and Recommendations for Best Practices. Couns Psychol. 2006;34(6):806-838. DOI: 10.1177/0011000006288127

52. Cortina JM. What is coefficient alpha? An examination of theory and applications. J Appl Psychol. 1993;78(1):98-104. DOI: 10.1037/0021-9010.78.1.98

53. Osborne J. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. Scotts Valley, CA: CreateSpace Independent Publishing; 2014.

54. Nunnally JC. Psychometric theory. 2nd ed. New York (NY): McGraw-Hill; 1978.

55. Kruyen PM, Emons WH, Sijtsma K. On the Shortcomings of Shortened Tests: A Literature Review. Int J Testing. 2013;13(3):223-248. DOI: 10.1080/15305058.2012.703734

56. Crowley SL, Fan X. Structural equation modeling: basic concepts and applications in personality assessment research. J Pers Assess. 1997;68(3):508-531. DOI: 10.1207/s15327752jpa6803_4

57. Hooper D, Coughlan J, Mullen M. Structural Equation Modeling: Guidelines for Determining Model Fit. Elect J Business Res Meth. 2007;6.

58. Kline RB. Principles and practice of structural equation modeling. 4th ed. New York (NY): Guilford Press; 2016.

59. Kline RB. Principles and practice of structural equation modeling. New York (NY): Guilford Press; 1998.

60. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equat Mod. 1999;6(1):1-55. DOI: 10.1080/10705519909540118

61. Browne MW, Cudeck R. Alternative Ways of Assessing Model Fit. Sociol Methods Res. 1992;21(2):230-258. DOI: 10.1177/0049124192021002005

62. Jöreskog KG, Sörbom D. LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language. Scientific Software International; 1993.

63. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods. 1996;1(2):130-149.

64. Knekta E, Runyon C, Eddy S. One Size Doesn't Fit All: Using Factor Analysis to Gather Validity Evidence When Using Surveys in Your Research. CBE Life Sci Educ. 2019;18(1):rm1. DOI: 10.1187/cbe.18-04-0064

65. Martin CR, Savage-McGlynn E. A 'good practice' guide for the reporting of design and analysis for psychometric evaluation. J Reprod Infant Psychol. 2013;31(5):449-455. DOI: 10.1080/02646838.2013.835036

66. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. Psychol Methods. 1999;4(3):272-299. DOI: 10.1037/1082-989x.4.3.272

67. Comrey AL, Lee HB. A first course in factor analysis. 2nd ed. Hillsdale (NJ): L. Erlbaum Associates; 1992.

68. Kaiser HF. The Application of Electronic Computers to Factor Analysis. Educ Psychol Meas. 1960;20(1):141-151. DOI: 10.1177/001316446002000116

69. Guttman L. Some Necessary Conditions for Common-Factor Analysis. Psychometrika. 2025;19(2):149-161. DOI: 10.1007/bf02289162

70. Arbaugh JB, Cleveland-Innes M, Diaz SR, Garrison DR, Ice P, Richardson JC, Swan KP. Developing a community of inquiry instrument: Testing a measure of the Community of Inquiry framework using a multi-institutional sample. Internet High Educ. 2008;11(3):133-136. DOI: 10.1016/j.iheduc.2008.06.003

71. Zhang YY, Lin X, Xu M. Rovai's Classroom Community Scale and Its Application in Chinese College English Class. Sino-US Engl Teach. 2011;8(9):592-598.

72. Lowenthal PR, Trespalacios J. Classroom Community and Time: Comparing Student Perceptions in Traditional vs. Accelerated Online Courses. Online Learn. 2022;26(4):59-77. DOI: 10.24059/olj.v26i4.3498

73. Trespalacios J, Lowenthal P. What do they really like? An investigation of students' perceptions of their coursework in a fully online educational technology program. Australas J Educ Technol. 2019;35(5):60-78. DOI: 10.14742/ajet.4364

74. Lowenthal PR, Dunlap JC. Investigating students' perceptions of instructional strategies to establish social presence. Distance Educ. 2018;39(3):281-298. DOI: 10.1080/01587919.2018.1476844

75. Picton A, Greenfield S, Parry J. Why do students struggle in their first year of medical school? A qualitative study of student voices. BMC Med Educ. 2022;22(1):100. DOI: 10.1186/s12909-022-03158-4

76. Traus A, Höffken K, Thomas S, Mangold K, Schröer W. Stu.diCo. – Studieren digital in Zeiten von Corona. Hildesheim: UVH - Universitätsverlag Hildesheim; 2020.

77. Beeson E, Aideyan B, Shoney CO, Bowes DA, Ansell KL, Peterson HM. Predicting Sense of Community among Graduate Students in a Distance Learning Environment. Univers J Educ Res. 2019;7(3):746-753. DOI: 10.13189/ujer.2019.070314

78. Floyd FJ, Widaman KF. Factor analysis in the development and refinement of clinical assessment instruments. Psychol Assess. 1995;7(3):286-299. DOI: 10.1037/1040-3590.7.3.286

79. Olaru G, Witthöft M, Wilhelm O. Methods matter: Testing competing models for designing short-scale Big-Five assessments. J Res Person. 2015;59:56-68. DOI: 10.1016/j.jrp.2015.09.001

80. Schroeders U, Wilhelm O, Olaru G. The influence of item sampling on sex differences in knowledge tests. Intelligence. 2016;58:22-32. DOI: 10.1016/j.intell.2016.06.003

81. MacCallum R, Widaman K, Zhang S, Hong S. Sample Size in Factor Analysis. Psychol Methods. 1999;4:84-99. DOI: 10.1037/1082-989X.4.1.84

82. Kenny DA, McCoach DB. Effect of the Number of Variables on Measures of Fit in Structural Equation Modeling. Struct Equ Modeling. 2009;10(3):333-351. DOI: 10.1207/s15328007sem1003_1

83. Jackson DL, Gillaspy JA, Purc-Stephenson R. Reporting practices in confirmatory factor analysis: an overview and some recommendations. Psychol Methods. 2009;14(1):6-23. DOI: 10.1037/a0014694

84. Knof H, Huber V, Shiozawa T. Validierung einer Kurzform der Deutschen Version der Classroom Community Scale (CCS-D-SF). In: Jahrestagung der Gesellschaft für Medizinische Ausbildung (GMA). Dresden, 17.-19.09.2026. Düsseldorf: German Medical Science GMS Publishing House; 2026. DocV-0701. DOI: 10.3205/26gma037

Korrespondenzadresse:
Dr. med. Thomas Shiozawa
Eberhard Karls Universität Tübingen, Medizinische Fakultät, Institut für Klinische Anatomie und Zellanalytik, Elfriede-Aulhorn-Str. 8, 72076 Tübingen, Deutschland
thomas.shiozawa@uni-tuebingen.de

Bitte zitieren als
Knof H, Huber V, Shiozawa T. Validation of a short form of the German form of the classroom community scale (CCS-D-SF). GMS J Med Educ. 2026;43(7):Doc97.
DOI: 10.3205/zma001891, URN: urn:nbn:de:0183-zma0018918

Artikel online frei zugänglich unter
https://doi.org/10.3205/zma001891

Eingereicht: *01.11.2025*
Überarbeitet: *17.07.2026*
Angenommen: *10.08.2026*
Veröffentlicht: *16.11.2026*

Copyright
©2026 Knof et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe http://creativecommons.org/licenses/by/4.0/.