

An Investigation of the Pedagogical Content Knowledge across German Preservice (Physical Education) Teachers

Stefan Meier

Centre for Sport Science and University Sports & Centre for Teacher Education, University of Vienna, Vienna, Austria

Email: stefan.meier@univie.ac.at

How to cite this paper: Meier, S. (2021). An Investigation of the Pedagogical Content Knowledge across German Preservice (Physical Education) Teachers. *Advances in Physical Education*, 11, 340-352. <https://doi.org/10.4236/ape.2021.113029>

Received: June 6, 2021

Accepted: August 2, 2021

Published: August 5, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Pedagogical content knowledge (PCK) is of critical importance to Physical Education (PE), since teaching PE is fundamentally distinct from teaching other subjects in many significant ways. Despite the importance of PCK, research on PCK in German speaking countries is still at the beginning. Against this backdrop, the current study explores the extent to which PCK is a specific professional feature across German students aiming for a teaching degree in PE or not. A cross-sectional study was conducted among 762 students to explore potential differences in relation to teacher education (TE) programs (PETE students $n = 431$, TE students $n = 331$). Measurement invariance (MI) between the groups was carried out using multigroup confirmatory factor analysis models to ensure latent mean scores can be compared meaningfully. The progressive evaluation of MI confirms that it is possible to measure the PCK (scalar) equivalently across PETE and TE students. PETE students outperformed TE students in both PCK subdimensions, also in different stages of the study. The study provides evidence for the “professional knowledge” and “qualification hypothesis” within PETE programs.

Keywords

Measurement Invariance, Pedagogical Content Knowledge, Preservice Physical Education Teachers, Professional Knowledge

1. Introduction

There has been a growing interest in conducting research on teacher knowledge in recent decades. Following the influential work by Shulman (1986, 1987) researchers have been building on the concept of “pedagogical content knowledge” (PCK). Research on teaching and teacher education considers PCK to be a core

component of professional competence (Blömeke, Gustafsson, & Shavelson, 2015). Special attention has been directed toward teachers' PCK since it predicts both the quality of teaching as well as student learning (e.g., Baumert et al., 2010; Iserbyt et al., 2020). PCK is of critical importance, "since it deals with teachers' knowledge necessary to achieve the aims of teaching" (Depaepe et al., 2013: p. 15) by organizing, representing, and adapting content to the abilities and interests of learners and presented for instruction. Thus, PCK serves the function of providing the teacher with knowledge to transform the content in ways that make it understandable to students.

This is of special importance to Physical Education (PE), since teaching PE is fundamentally distinct from the teaching of other subjects in many significant ways. It can be named as almost the only subject which explicitly deals with body and corporeality on different levels (e.g., cognitive, socio-emotional). PE aims at fundamental experience with one's own body and connects this with reflection processes on one's own personal development, allows individual access to the body and thus to the world (Prohl, 2010). Thus, PE is the only compulsory subject whereby physical activity is a primary means of accomplishing educational objectives although with varying interpretations in different concepts across the European Union (EU) (MacPhail, Tannehill, & Avsar, 2019; Naul, 2003).

To date, research on PCK in the field of sport science has contained a "selection bias" (Depaepe et al., 2013: p. 22; Ward & Ayvazo, 2016: p. 201), because different didactics foci in the EU and research traditions on subject matter didactics (Van Driel & Berry, 2012) have received little interest. With particular respect to German speaking countries, research on PCK is still at the beginning (Vogler et al., 2018; Baumgartner, 2018; Heemsoth, 2016; Heemsoth & Wibowo, 2020; Vogler, Messmer, & Allemann, 2017; Wibowo & Heemsoth, 2019). As theoretical approaches to conceptualize PCK differ, German scholars mostly refer to dispositional orientated approaches (Vogler et al., 2018), whereas understandings of PCK "have been largely behavioural" in English-speaking publications (Backman & Barker, 2020: p. 2). Dispositional orientated approaches restrict the term competence to the sum of cognitive and motivational resources, assuming these multiple constituents are necessary for competent performance. Behavioral orientated approaches refer on how cognition, affect-motivation and performance are interlinked as a system and change during the in-situation performance (Blömeke, et al., 2015). A more integrated perspective focuses on the processes connecting both approaches (Krauss et al., 2020; for PE Baumgartner, 2018). However, as PCK is important in the sense of student learning, it is of special interest to explore the extent to which PCK is a specific professional feature, thus providing insights into the conditions of PCK within teacher education (TE) programs.

The highly specialized PCK is considered to be one of the main features distinguishing teachers from laypeople (Bromme, 2008; Mieg, 2001). Thus, PCK characterizes teachers' professional identity in a subject, also known as "professional knowledge hypothesis" (Baumert & Kunter, 2006; Krauss et al., 2008). For

instance, PE teachers are professionals in at least two areas: they are both professionals in sport science and professional teachers, whereas teacher students not aiming for a PE teaching degree are solely professional teachers. The latter are “related professionals” (Krauss et al., 2008: p. 881) as both PE and TE students have high levels of pedagogical expertise. To the best of our knowledge, there is no study investigating whether students aiming for a PE teaching degree or not differ in their PCK. With respect to other domains scholars reported higher PCK scores from students aiming for a subject specific teaching degree (Jüttner & Neuhaus, 2013; Krauss et al., 2008; Schmelzing et al., 2012). Hence, one could assume differences in the PCK between PETE and TE students.

In addition, scholars have evidenced that teacher education and professional development programs provided opportunities to acquire PCK (Richter, 2013), also known as “qualification hypothesis” or “growing knowledge hypothesis” (Krauss et al., 2008). As a consequence, beginner teachers adhered more to their written plan, while more experienced teachers were able to depart from their plan to provide PCK in accordance with their students’ abilities (Ward & Ayvazo, 2016). Such learning opportunities during teacher education and professional development programs have fostered PCK and in turn students’ learning (Iserbyt, Ward, & Martens, 2016; Iserbyt et al., 2020; Kim et al., 2018). In Germany a validation study showed that the PETE students’ semester predicted PCK, whereas the grade point average did not. This finding supports the qualification hypothesis (Heemsoth & Wibowo, 2020).

Against this background, the purpose of the current study is to compare the PCK of German students aiming for a teaching degree in PE or not. To date, no investigation has been made into this issue. Although the number of studies measuring PCK is rising (Meier, 2020, 2021; Heemsoth & Wibowo, 2020; Vogler et al., 2017), no scale has been tested for measurement invariance across these two groups. As PCK develops through different educational programs and other learning opportunities, measurement invariance is the precondition for comparing the PCK of such different groups. A meaningful and valid comparison of the PCK of both groups can be made only if a scale measures the same construct in both groups in the same way (Chen, 2008). The aims of the current study are as follows:

- 1) Is it possible to measure PCK equivalently across students aiming for a teaching degree in PE or not?
- 2) With regard to the “professional knowledge hypothesis”, we compare latent mean scores of PCK in students aiming for a teaching degree in PE or not. This approach is conservative as both students aim for a teaching degree, thus are related professionals. From this point of view, the professional knowledge hypothesis aims at analyzing to which extent PCK is deeply ingrained in the populations investigated. We hypothesize that PETE students score higher on PCK than TE students.
- 3) Towards the “qualification hypothesis”: Based on evidence that the PETE students’ semester predicted PCK, we hypothesize that PETE students score higher

on PCK than TE students in different stages of the study.

2. Materials and Methods

2.1. Study Design

A cross-sectional survey design was used to investigate the PCK across students aiming for a teaching degree in PE or not. The research was conducted in classes during regular courses. After receiving approval from the program directors, paper-and-pencil tests were brought to the courses. Surveys were conducted by trained test administrators as power tests without time limits. Participation was on a voluntary basis. The questionnaire included a covering letter with information about the purpose of the study, the benefits of participating in the study, and ethical issues related to anonymity and voluntariness. Basic information was collected on demographic variables (e.g., gender, age,). After the survey, participants could ask questions about the study in more detail. The students did not receive incentives or compensation for their participation.

2.2. Participants

As shown in **Table 1**, the global number of participants was 762 students in two different teacher education programs aged between 18 and 37. Most of the students were in their 2nd year of study. All participants were recruited from three public universities in one federal state of Germany, North-Rhine Westphalia. The sample comprises more males compared to females. One part of the sample comprised 431 PETE students aiming to become a teacher for upper secondary schools, the equivalent to International Standard Classification of Education (ISCED) 3. The other part of the sample comprised 331 teacher students not aiming for a PE teaching degree for ISCED 3. In Germany, teacher candidates decide at the very beginning of their study in which type of school they want to work after their graduation.

Table 1. Demographic characteristics of participants.

	M	N ^a	%
Age (SD)	22.62 (3.00)	757	
Semester (SD)	4.04 (2.74)		
	Beginner	387	51.67
	Advanced	362	48.33
Gender			
	Female	368	48.36
	Male	393	51.64
Study Program			
	PETE	431	55.56
	TE	331	43.44

a. Sample sizes vary due to missing data.

2.3. Measurements

To measure the PCK of the students, the 15-item “PCK-PE” (Meier, 2020) was used. The items covered two conceptually different PCK subscales: 1) knowledge of instructional strategies and 2) knowledge of students’ (mis)conceptions and difficulties. The “instruction” dimension stresses on different representations and explanations of making the content comprehensible to learners. The “students” dimension assessed the ability to recognize students’ (pre)conceptions about PE. The itemset consists of a mixture of open ended and multiple-choice questions. Responses were coded right or wrong by two trained raters following a standardized manual. The PCK-PE showed factorial and discriminant validity and good internal consistency for the subscales in prior research (Meier, 2020, 2021). In this study, the internal consistency of both the “instruction dimension” ($\alpha = 0.757$) and the “students’ dimension” ($\alpha = 0.815$) were good. The latent correlations between both PCK dimensions as computed on the basis of a configural invariance confirmatory factor analyses (CFA) model were 0.144 (PETE students) and 0.333 (TE students). Discrimination between the two constructs of PCK was therefore highest in the PETE students’ group.

2.4. Data Analysis

The data processing and frequency analyses were conducted using SPSS 26. Correlation and multivariate analyses were calculated with Mplus 8.3 (Muthén & Muthén, 2012).

The first research question addresses the comparison of latent means between the two different teaching degree programs (PETE or not) preparing for a PE specific teaching degree or another subject specific teaching degree. For this purpose, we investigated whether the testing instrument measured the constructs in the same way across these two groups, i.e., that the underlying constructs were invariant (equivalent) across different groups (Chen, 2008; van De Schoot, Schmidt, & De Beuckelaer, 2015). To test for measurement invariance (MI) between the groups, we therefore conducted a series of CFA following an approach that is well established in the literature on structural equation modeling. Based on pre-existing findings we tested the two-factor structure model of the “PCK-PE” through CFA including a review of modification indices. After that we conducted a CFA to compare the fit of this two-factor model with the G-factor model to figure out the most parsimonious model for the following MI analyses. Following that, several nested models of multigroup CFAs (MGCFA) were conducted to study MI within the framework of structural equation modelling to determine the extent to which the factor structure was comparable across the study program and stages of the study. This approach involves setting cross-group constraints on parameters and comparing more restricted models with less restricted models (Millsap, 2011). For the MI of categorically ordered data three steps were considered (Muthén & Muthén, 2012): the baseline model tested the original two-factor structure through a CFA for each group separately. Pro-

ceeding to the simultaneous MGCFA, the configural model and the scalar model were tested, from least restrictive to most restrictive. More restricted models were compared with less restricted models. Scalar equivalence (strong MI) across groups is a precondition for comparing means.

For all analyses, the means and variance adjusted weighted least squares estimator (WLSMV) was chosen because the data are categorical (Flora & Curran, 2004). The Mplus DIFFTEST option was used to perform χ^2 difference tests for the nested model comparison evaluation. As chi-square tests (χ^2) are sensitive to sample size and may reject models with even trivial misfit (Chen, 2007), we used the root mean square error of approximation (RMSEA, cut-off value for a good model fit <0.06 , acceptable fit <0.08) and the comparative fit index (CFI, cut-off value for a good fit was >0.95 and acceptable fit >0.90) to evaluate goodness of fit (Hu & Bentler, 1999; Marsh, Hau, & Wen, 2014). Chi-square difference tests between the nested models were applied in which the difference in χ^2 value ($\Delta\chi^2$) relative to the change in degrees of freedom (Δdf) was evaluated, as were changes in RMSEA ($\Delta RMSEA$) and CFI (ΔCFI). Model equivalence was indicated by either a nonsignificant $\Delta\chi^2$ or ΔCFI values ≤ 0.010 and $\Delta RMSEA$ values ≤ 0.015 (Chen, 2007; Cheung & Rensvold, 2002; Rutkowski & Svetina, 2017).

To address the second and third research question (“professional knowledge” and “qualification hypothesis”), we examined differences in latent means across the study program and stages of the study as these have been reported as significant to PCK (Heemsoth & Wibowo, 2020; Iserbyt et al., 2020; Ward & Ayvazo, 2016). In Mplus effect sizes are not directly computed, so to determine the magnitude of differences in latent means, we calculated an effect size d for these differences. Common standards for small, medium, and large standardized effects are 0.2, 0.5, and 0.8, respectively (Cohen, 1988).

3. Results

3.1. Factor Structure of the PCK-PE

Based on prior findings, we hypothesized that a two-factor model would be an appropriate fit with the data (Meier, 2020, 2021). This model differentiates the two latent dimensions “instruction” and “students”. The CFA of this initial two-factor model resulted in an acceptable model fit: $\chi^2 (df) = 415.752 (89)$, $p < 0.001$, CFI = 0.976, RMSEA = 0.069. To find a (more) parsimonious and well-fitting model, we reviewed modification indices. Although there were a few additional modifications, we did not make additional changes since it did not result in significant changes in fit indices. Subsequently the G-factor model CFA was carried out on the basis of the initial two-factor model and resulted in a worse fit to the data, with all indices being worse compared to the initial model. The $\Delta\chi^2$ result indicated that the initial two-factor model fitted the data significantly better than the G-factor model ($\Delta\chi^2 (\Delta df) = 224.108 (1)$, $p < 0.001$). Before the MI analysis, the initial two-factor model was tested on different study program and stages of study. Indices revealed that the two-factor model gener-

ally fits the data well in each subsample. Thus, the two-factor model can serve as the initial model for the subsequent MI tests.

3.2. Measurement Invariance across Students

The fit indices for the basic model in subsamples and each MI test step are shown in **Table 2**. Since there were two indicators loading on two factors, only the configural invariance model and scalar invariance model were tested in each group (i.e., study program, stages of the study). First, we investigated MI across study program, in which students were classed into a PETE group and TE group. This distinction reflects different paths of subject matter education in the context of teacher training and is in line with the “professional knowledge hypothesis”. Results of the configural and scalar invariance model indicated that the two-factor structure was verified across study program. Both models fitted the data well. Though the χ^2 difference test showed a significant χ^2 (df) change, the changes in CFI and RMSEA values from the invariance configural model showed that the constrained model was not rejected. According to the “qualification hypothesis” (e.g., Heemsoth & Wibowo; Iserbyt et al., 2020), we investigated whether stages of study affected the measurement model. Based on the year of study (self-reported), students were split into a beginner (1st year students) and a more advanced group (2nd year students and older). The χ^2 difference test suggested that there is no significant deterioration in the model fit between the configural and scalar invariance model. In addition, the increase in CFI and RMSEA indicated an equal fit. Since all MI tests provided evidence for configural and scalar invariance of the two-factor PCK model in the PE(TE) students group and subsamples, comparisons of latent group-mean PCK scores seemed to be acceptable.

Table 2. Measurement invariance model fitting indices and comparison across the two groups (PETE and TE students).

Model ^a	χ^2	df	<i>p</i>	RMSEA	CFI	Δ RMSEA	Δ CFI	$\Delta\chi^2$ (Δ df)	<i>p</i>
Study Program									
PETE (n = 431)	222.544	89	0.000	0.059	0.981				
TE (n = 331)	199.520	89	0.000	0.061	0.974				
Configural	420.935	178	0.000	0.060	0.978				
Scalar	452.812	189	0.000	0.061	0.976	0.001	0.002	37.922 (11)	0.0001
Stage of the study ^b									
Beginner (n = 387)	222.972	89	0.000	0.062	0.977				
Advanced (n = 362)	277.990	89	0.000	0.077	0.980				
Configural	503.263	178	0.000	0.070	0.979				
Scalar	512.915	189	0.000	0.068	0.979	0.002	0.000	8.577 (11)	0.6609

a. $\Delta\chi^2$ test is conducted using the DIFFTEST option for nested models. b. 13 missing values.

3.3. Latent Mean Differences in PCK

The differences between groups in the latent means for the two constructs of the PCK-PE are shown in **Table 3**. TE students scored significantly lower compared to the PETE students in both, the “instruction dimension” ($d = 0.883$) as well as in the “students’ dimension” ($d = 0.984$). The effect sizes for the differences are large. By comparing the means in both the study programs in different stages of their study separately, TE students scored lower on both dimensions than PETE students at the beginning and at the end of studying. The effect size for the difference at the beginning was medium in the “instruction dimension” ($d = 0.661$) and large in the “students’ dimension” ($d = 0.984$). At the end of studying it was large in both the “instruction dimension” ($d = 1.090$) and in the “students dimension” ($d = 1.041$).

Table 3. Comparison of Latent Means in PCK-PE Subscales via Series of MGCFA.

	PCK-PE Instruction			PCK-PE Students		
	M	SE	<i>P</i>	M	SE	<i>P</i>
Study Program						
PETE (n = 431)	0.000	0.000		0.000	0.000	
TE (n = 331)	-1.110	0.180	0.000	-1.080	0.081	0.000
Stage of the study*/Study Program						
Beginner/Study Program (n = 387)						
PETE (n = 203)	0.000	0.000		0.000	0.000	
TE (n = 184)	-0.669	0.159	0.000	-1.054	0.106	0.000
Advanced/Study Program (n = 362)						
PETE (n = 225)	0.000	0.000		0.000	0.000	
Sport Science (n = 137)	-1.942	0.553	0.000	-1.137	0.132	0.000

a. 13 missing values.

4. Discussion

The purpose of the current study was the PCK of German students aiming for a teaching degree in PE or not. As research on PCK in German-speaking countries is still at the beginning, this research contributes to a more comprehensive picture of the PCK, accounting for different didactics foci and research traditions on subject matter didactics in the EU. The aim was to examine the “professional knowledge” and the “qualification hypothesis” within (PE)TE programs.

The first research question tested the factor structure of the PCK-PE across students aiming for a teaching degree in PE or not via CFA. With the MGCFA procedure, we ensured that the factor structure was invariant across groups. In order to make a reliable comparison of the PCK-PE scores between the two groups, examination of MI is fundamental. The MI analysis indicated that the conceptual framework to define the two latent factors (the “instruction” and “students’ dimension”) is equivalent for PETE and TE students and in different

stages of the study (beginner vs. advanced). Hence, it makes sense to compare the mean scores between aiming for a teaching degree in PE or not in these different conditions (Chen, 2008; Cheung & Rensvold, 2002). This is of critical importance as, to date, there is no evidence towards MI across such groups.

With the second research question, the subsequent latent mean comparisons between students aiming for a teaching degree in PE or not provide further insights into the “professional knowledge hypothesis”. As hypothesized, the PCK is a special feature distinguishing PETE-students’ from TE students in other subjects. PETE students outperform the TE students in both PCK-PE subdimensions the “instruction” and “students’ dimension”. The effect sizes for the differences in both PCK-PE subdimensions were large. This finding is consistent with findings in other domains (Jüttner & Neuhaus, 2013; Schmelzing et al., 2012) supporting the assumption that PCK characterizes a teachers’ professional identity in a subject. PETE students aim to be professional on at least two dimensions: sport science and teaching. TE students aim to become a teacher in other subjects and thus, they are related professionals. With regard to the “professional knowledge hypothesis” PCK distinguishes PETE students from TE students. This finding sensitizes for a deeper understanding of the PCK-difference, which could date back to learning opportunities in PETE: To what extent foster (which kind of) learning opportunities during PETE the PCK? However, future research should study this in more detail. Furthermore, the study stressed on PCK in the sense of dispositional oriented approaches. As a result, it remains an open question, how PCK is linked to observable performance. In line with Krauss et al. (2020: p. 312) future research should “trace the long route from teacher disposition to student learning” and thus examine how much PCK predicts student learning.

With respect to the third research question, mean comparison in different stages of the study highlighted that TE students scored significantly lower than PETE students on both the “instruction” and “students’ dimension” at the beginning as well as at the end of studying. The effect sizes for these differences were medium to large, especially at the end of studying. As we drew on cohort comparisons (beginner vs. advanced) and not on large scale data, this provides little evidence for the “qualification hypothesis”: learning opportunities during PETE are conducive to the development of the highly specialized PCK, which is in line with prior findings (e.g., Heemsoth & Wibowo, 2020; Iserbyt et al., 2020). However, this must be tackled in more detail in future studies.

5. Conclusion

In this study we measured the PCK of German students aiming for a teaching degree in PE or not and investigated differences in latent means in subsamples (i.e., stages of the study). The results provided evidence that the factor structure of the PCK was invariant across (sub)groups, thus latent mean scores can be compared meaningfully. In line with the “professional knowledge hypothesis”,

PETE students outperformed TE students in both PCK subdimensions, which is also pertinent in different stages of the study. With regard to the “qualification hypothesis” one could argue that study progress within PETE program fosters the development of PCK as TE students scored significantly lower on both PCK subdimensions at the beginning as well as at the end of studying.

However, when interpreting the results of the present study some limitations need to be considered. Firstly, participants were all from one region in Germany, aiming for a specific teaching degree (i.e., ISCED 3) in PE and other subjects, therefore the results apply only to these study programs, preventing generalizations. Secondly, the complex nature of cross-sectional design prevents us from drawing causal conclusions. Although we tested MI and considered covariates, we cannot rule out that other factors confound the group differences in PCK. Remarkably, the cohort-comparison of students in different stages (beginner vs. advanced) must be taken as a tendency. Given the said limitations of this study, these observations should be treated with caution. Future, preferably longitudinal studies with prospective investigation, should study the extent to which PCK develops during PETE programs and control for covariates (e.g., learning opportunities). Furthermore, it must be pointed out that it is an open question to what degree, the findings of the current study relate to observable teacher behavior in class and thus future studies should relate the measure of PCK with teacher performance. Finally, the evaluation of (MG)CFA models and MI with categorical indicators is a field not well studied. Although the number of studies is rising, recommendations for using fit measures and cut-off values are based on only a few simulation studies (Rutkowski & Svetina, 2017).

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- Backman, E., & Barker, D. M. (2020). Re-Thinking Pedagogical Content Knowledge for Physical Education Teachers—Implications for Physical Education Teacher Education. *Physical Education and Sport Pedagogy*, *25*, 451-463. <https://doi.org/10.1080/17408989.2020.1734554>
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, *9*, 469-520. <https://doi.org/10.1007/s11618-006-0165-2>
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Tsai, Y.-M. et al. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal*, *47*, 133-180. <https://doi.org/10.3102/0002831209345157>
- Baumgartner, M. (2018). “Kompetenz ohne Performanz ist leer! Performanz ohne Kompetenz blind...!” Zu einem integrativen Kompetenzstrukturmodell von Sportlehrkräften. *Zeitschrift für sportpädagogische Forschung*, *6*, 49-68.
- Blömeke, S., Gustafsson, J. E., & Shavelson, R. J. (2015). Beyond Dichotomies: Competence

- Viewed as a Continuum. *Zeitschrift für Psychologie*, 223, 3-13.
<https://doi.org/10.1027/2151-2604/a000194>
- Bromme, R. (2008). Lehrerexpertise. In W. Schneider, & M. Hasselhorn (Eds.), *Handbuch der Pädagogischen Psychologie* (pp. 159-167). Hogrefe.
- Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling*, 14, 464-504.
<https://doi.org/10.1080/10705510701301834>
- Chen, F. F. (2008). What Happens If We Compare Chopsticks with Forks? The Impact of Making Inappropriate Comparisons in Cross-Cultural Research. *Journal of Personality and Social Psychology*, 95, 1005-1018. <https://doi.org/10.1037/a0013193>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling*, 9, 233-255.
https://doi.org/10.1207/S15328007SEM0902_5
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Erlbaum Associates.
- Depaepe, F., Verschaffel, L., & Kelchtermans, G. (2013). Pedagogical Content Knowledge: A Systematic Review of the Way in Which the Concept Has Pervaded Mathematics Educational Research. *Teaching and Teacher Education*, 34, 12-25.
<https://doi.org/10.1016/j.tate.2013.03.001>
- Flora, D. B., & Curran, P. J. (2004). An Empirical Evaluation of Alternative Methods of Estimation for Confirmatory Factor Analysis with Ordinal Data. *Psychological Methods*, 9, 466-491. <https://doi.org/10.1037/1082-989X.9.4.466>
- Heemsoth, T. (2016). Fachspezifisches Wissen von Sportlehrkräften. Ein Überblick über fachübergreifende und fachfremde Ansätze und Perspektiven für die Professionsforschung von Sportlehrkräften. *Zeitschrift für sportpädagogische Forschung*, 4, 41-60.
- Heemsoth, T., & Wibowo, J. (2020). Fachdidaktisches Wissen von angehenden Sportlehrkräften messen. *German Journal of Exercise and Sport Research*, 50, 308-319.
<https://doi.org/10.1007/s12662-020-00643-0>
- Hu, L., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling*, 6, 1-55. <https://doi.org/10.1080/10705519909540118>
- Iserbyt, P., Coolkens, R., Looockx, J., Vanluyten, K., Martens, J., & Ward, P. (2020). Task Adaptations as a Function of Content Knowledge: A Functional Analysis. *Research Quarterly for Exercise and Sport*, 91, 539-550.
<https://doi.org/10.1080/02701367.2019.1687809>
- Iserbyt, P., Ward, P., & Martens, J. (2016). The Influence of Content Knowledge on Teaching and Learning in Traditional and Sport Education Contexts: An Exploratory Study. *Physical Education and Sport Pedagogy*, 21, 539-556.
<https://doi.org/10.1080/17408989.2015.1050662>
- Jüttner, M., & Neuhaus, B. J. (2013). Das Professionswissen von Biologielehrkräften. Ein Vergleich zwischen Biologielehrkräften, Biologen und Pädagogen. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 31-49.
- Kim, I., Ward, P., Sinelnikov, O., Ko, B., Iserbyt, P., Li, W., & Curtner-Smith, M. (2018). The Influence of Content Knowledge on Pedagogical Content Knowledge: An Evidence-Based Practice for Physical Education. *Journal of Teaching in Physical Education*, 37, 133-143. <https://doi.org/10.1123/jtpe.2017-0168>
- Krauss, S., Baumert, J., & Blum, W. (2008). Secondary Mathematics Teachers' Pedagogical Content Knowledge and Content Knowledge: Validation of the COACTIV Constructs. *ZDM: The International Journal on Mathematics Education*, 40, 873-892.

- <https://doi.org/10.1007/s11858-008-0141-9>
- Krauss, S., Bruckmaier, G., Lindl, A., Hilbert, S., Binder, K., Steib, N., & Blum, W. (2020). Competence as a Continuum in the COACTIV Study: The “Cascade Model”. *ZDM: The International Journal on Mathematics Education*, 52, 311-327. <https://doi.org/10.1007/s11858-020-01151-z>
- MacPhail, A., Tannehill, D., & Avsar, Z. (2019). *European Physical Education Teacher Education Practices. Initial, Induction, and Professional Development*. Meyer & Meyer.
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler’s (1999) Findings. *Structural Equation Modeling*, 11, 320-341. https://doi.org/10.1207/s15328007sem1103_2
- Meier, S. (2020). Development and Validation of a Testing Instrument to Assess Pedagogical Content Knowledge of German Preservice Physical Education Teachers. *Journal of Physical Education and Sport*, 20, 3010-3016.
- Meier, S. (2021). Pedagogical Content Knowledge in Students Majoring in Physical Education vs. Sport Science. The Same but Different? *German Journal of Exercise and Sport Research*. <https://doi.org/10.1007/s12662-021-00725-7>
- Mieg, H. (2001). *The Social Psychology of Expertise: Case Studies in Research, Professional Domains, and Expert Roles*. Lawrence Erlbaum. <https://doi.org/10.4324/9781410602282>
- Millsap, R. E. (2011). *Statistical Approaches to Measurement Invariance*. Routledge. <https://doi.org/10.4324/9780203821961>
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus User’s Guide*. Muthén & Muthén.
- Naul, R. (2003). Concepts of Physical Education in Europe. In K. Hardman (Ed.), *Physical Education: Deconstruction and Reconstruction, Issues and Directions* (pp. 35-52). Hofmann.
- Prohl, R. (2010). *Grundriss der Sportpädagogik*. Limpert.
- Richter, D. (2013). Professional Development across the Teaching Career. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers Results from the COACTIV Project* (pp. 333-342). Springer.
- Rutkowski, L., & Svetina, D. (2017). Measurement Invariance in International Surveys: Categorical Indicators and Fit Measure Performance. *Applied Measurement in Education*, 30, 39-51. <https://doi.org/10.1080/08957347.2016.1243540>
- Schmelzing, S., van Driel, J., Jüttner, M., Brandenbusch, S., Sandmann, A., & Neuhaus, B. J. (2012). Development, Evaluation, and Validation of a Paper-and-Pencil Test for Measuring Two Components of Biology Teachers’ Pedagogical Content Knowledge Concerning the ‘Cardiovascular System’. *International Journal of Science and Mathematics Education*, 11, 1369-1390. <https://doi.org/10.1007/s10763-012-9384-6>
- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 15, 4-14. <https://doi.org/10.3102/0013189X015002004>
- Shulman, L. S. (1987). Knowledge and Teaching. Foundations of the New Reform. *Harvard Educational Review*, 57, 1-22. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijenburg, M. (2015). Editorial: Measurement Invariance. *Frontiers in Psychology*, 6, Article ID: 1064. <https://doi.org/10.3389/fpsyg.2015.01064>
- Van Driel, J. H., & Berry, A. (2012). Teacher Professional Development Focusing on Pe-

dagogical Content Knowledge. *Educational Researcher*, 41, 26-28.

<https://doi.org/10.3102/0013189X11431010>

Vogler, J., Messmer, R., & Allemann, D. (2017). Das Fachdidaktische Wissen und Können von Sportlehrpersonen (PCK-Sport). *German Journal of Exercise and Sport Research*, 47, 335-347. <https://doi.org/10.1007/s12662-017-0461-4>

Vogler, J., Messmer, R., Wibowo, J., Heemsoth, T., & Meier, S. (2018). Drei Zugänge zur Modellierung fachdidaktischen Wissens von Sportlehrpersonen. In E. Balz, & D. Kuhlmann (Eds.), *Sportwissenschaft in pädagogischem Interesse* (pp. 47-55). Czwalina. <https://doi.org/10.1007/s12662-017-0461-4>

Ward, P., & Ayvazo, S. (2016). Pedagogical Content Knowledge: Conceptions and Findings in Physical Education. *Journal of Teaching in Physical Education*, 35, 194-207. <https://doi.org/10.1123/jtpe.2016-0037>

Wibowo, J., & Heemsoth, T. (2019). Zur Konzeption und Validität eines Tests zum fachdidaktischen Wissen von Sportlehrkräften. *Zeitschrift für sportpädagogische Forschung*, 7, 88-108.