Research Article

Journal of Educational & Psychological Research

Measuring Teaching Quality and Student Engagement in Elementary Education in South Korea and The Netherlands

Wim van de Grift^{1*}, Seyeoung Chun² and Okhwa Lee³

¹Van de Grift: Research and Advice, Soestduinen, The Netherlands

²Chungnam National University, Deajeon, South Korea

³Chungbuk National University, Cheongju, South Korea

*Corresponding Author

Wim van de Grift, Research and Advice, Soestduinen, The Netherlands.

Submitted: 2023, Oct 23; Accepted: 2023, Nov 14; Published: 2023, Nov 24

Citation: Van de Grift, W., Chun, S., Lee, O. (2023). Measuring Teaching Quality and Student Engagement in Elementary Education in South Korea and The Netherlands. *J Edu Psyc Res*, 5(3), 767-772.

Abstract

254 South Korean and 384 Dutch teachers working in elementary education were observed by specially trained observers. The observation-scales are sufficiently homogeneous and valid for estimating sample averages. Inter rater agreement shows moderate/good mutual consensus. Results of multigroup confirmatory factor analysis show that comparing correlations and average scores is allowed for both countries. Engagement of students is a bit better in South Korea than in the Netherlands. South Korean teachers do not differ from Dutch teachers on the efficiency of their classroom management and on giving clear instruction. South Korean teachers are a bit better than Dutch teachers on activating students and clearly better on teaching learning strategies. Dutch teachers are a bit better on creating a safe educational climate and on differentiation of instruction.

Keywords: Effective Teaching, Observation, International Comparative Analysis

Highlights

South Korean students are a bit better engaged in the lessons, than Dutch students.

South Korean teachers do not differ from Dutch teachers in:

- organizing efficient classroom management, and
- giving clear instruction.

South Korean teachers are bit better than Dutch teachers in:

- activating students, and
- clearly better teaching learning strategies.

Dutch teachers are a bit better than South Korean teachers in:

- creating a safe learning climate, and
- differentiation of instruction.

1. Introduction

According to studies of the Programme for International Student Assessment (PISA) of the Organisation for Economic Co-operation and Development (OECD), published in 2012, 2014 and 2018, South Korean 15-year-olds scored on average 23-60% of a standard deviation higher than the OECD average on reading, mathematics and science. According to Cohen and

Lipsey these differences may be qualified as small to medium size differences [1, 2]. The average scores of Dutch students for reading, mathematics, and science are in these years 7-31% of a standard deviation lower than the average scores of South Korean students. These differences were reason for starting several comparative studies into differences in the quality of the teaching of the teachers in elementary and secondary education in both countries. A study in secondary education revealed that 289 Dutch teachers and 375 South Korean teachers in secondary education did not differ in the quality of 'clear and structured instruction'. The Dutch teachers scored significantly better on basic skills like 'creating a safe and stimulating learning climate' and 'intensive and activating teaching' and almost significantly better on 'efficient classroom management'. The South Korean teachers performed significantly better than the Dutch teachers on the advanced skills 'teaching learning strategies' and almost significantly on 'differentiating instruction'. Furthermore, a significant better student engagement was found in South Korea [3].

Teaching Skill	SK (n=375)	sd	NL (n=289)	sd	Cohen's δ	sign.		
safe learning climate	3.04	.67	3.27	.55	.37	.000		
classroom management	3.05	.69	3.16	.51	.18	.052		
clear instruction	2.91	.64	2.97	.52	.10	.155		
activating students	2.69	.67	2.82	.56	.21	.009		
differentiation	2.36	.83	2.26	.46	14	.054		
teaching learning strategies	2.57	.73	2.45	.53	18	.024		
student engagement	3.09	.70	2.91	.62	27	.001		
(cf. Van de Grift, Chun, Maulana, Lee, Helms-Lorenz, 2017) [3]								

Table 1: Average Scores of South Korean and Dutch Teachers in Secondary Education

The next question is: Do these differences in the quality of teaching already exist in elementary education? Or more precisely: What are the differences in teaching skills and student engagement in elementary education in South Korea and the Netherlands.

2. Theoretical and Empirical Background

This study finds its theoretical and empirical grounds in an international research program that started around the seventies, after the disappointing conclusions about the impact of teachers on the performance of American students [4]. The results of this 50 years during research program are summarised by Hattie; Marzano; Muijs & Reynolds; Scheerens and others [5-9]. One of the most important results of this international research program is that that student's achievements are not only caused by the students' intelligence and socio-economic background; the teacher too matters in determining student's achievements. A lot of studies made clear that about 15-25% of differences in students' achievement might be explained by the work of teachers [10-32].

Several studies even showed that an average growth in the skill of teachers with half a standard deviation goes along with an average improvement of 20% of a standard deviation of the performance of their students [18-20, 33, 34].

Educational economists, working with secondary analyses of original educational effectiveness studies, conclude that students of better skilled teachers, may earn annually on average \$20,000 more, later on [16, 17].

3. Aim of this Study

This study is aimed at comparing the differences of the quality of teaching and student engagement in elementary education in South Korea and the Netherlands.

Such a comparison offers opportunities to study whether the quality of teaching and learning is a more or less universal or a more or less culturally bound issue. There are better opportunities for improving education, if the quality of learning and teaching is a more or less universal matter, than when cultural differences play too big a role. Therefore, it is very important that we are able to measure teaching skills and student engagement in both countries in a reliable and valid way and we that we are sure that in both countries the factor loadings and intercepts are equal for

each of the observation scales. That is why we not only study the classical psychometric properties of the observation instrument, but also the possibilities to compare the correlations and the average scores in both countries despite cultural differences.

This study is aimed at answering the following more specific questions:

- Are we able to measure teaching skills and student engagement in both countries in a reliable and valid way?
- Are we allowed to compare correlations and average scores in both countries?
- What are the differences in teaching skills and student engagement in both countries?

4. Method

4.1. Samples

In 2016, 254 teachers working in elementary education were observed in the Daejeon and Cheongju region in South Korea. In 2016, the South Korean population of schools for primary education consisted of 6280 schools and 163645 teachers.

In the Netherlands, during 384 teachers working in elementary education with 6-12 year old students were observed. In 2016, the Dutch population of schools for primary education consisted of 6508 schools and about 154000 teachers. For these population sizes, samples of 385 or more observations are needed to achieve $\pm 5\%$ accuracy in calculating the population mean with a 95% confidence level. Samples of 262 are large enough to find significant (α <.05 with a power of .80) differences with effect sizes (Cohen's δ) of about .20 and more.

4.2. Observation Instrument

The ICALT-observation-instrument was used for observing the teachers in both countries. The items of the ICALT-observation-instrument are based on the results of several educational effectiveness studies mentioned in section 2. From these studies a total of 32 high inferential observable teaching activities and 120 low inferential observable teaching activities were obtained [35, 36]. These 152 activities are arranged into six domains: safe and stimulating educational climate, efficient classroom management, clear and structured instruction, intensive and activating teaching, differentiating instruction, and teaching learning strategies. The items have been used for the first time by the Dutch inspectorate of education in order to evaluate the quality of teaching in primary education in the Netherlands

[35, 37]. Several studies have been done into the reliability and validity of this instrument in elementary education in the Netherlands [36, 38-40]. Various versions of the ICALT observation instrument can be found in Van de Grift, 2007; Van de Grift et al, 2019. For the most recent version, send an email to: Wim.vandeGrift@ZIGGO.NL [36, 41].

4.3. Training of Observers and Inter Rater Reliability

All participating observers in both countries attended a training of half a day. They got information about the theoretical and empirical backgrounds of the ICALT-observation-instrument. During the training the observers watched and scored two video recorded lessons. The results of the first video were used for discussions between the observers in order to find agreement between observers. In both countries, the results of the observations of the second video recorded lesson were used to test whether the mutual agreement reached at least a moderate/good mutual consensus (Fleiss's κ of >.60).

4.4. Homogeneity of the Scales

For both South Korea and the Netherlands, we computed the classical homogeneity coefficient Cronbach's α [42]. Norm for acceptable reliability is: \geq .70. The results are found in table 2.

basic skills			advanced ski	ills	all 32	student		
	safe learning climate	classroom management	clear instruction	activating students	differentiation	teaching learning strategies	ICALT items	engagement
South Korea	.80	.81	.82	.83	.78	.87	.95	.83
Netherlands	.85	.85	.82	.80	.79	.87	.93	.81

Table 2: Homogeneity (Cronbach's α) of the Likert Scales

All scales are in both countries sufficiently homogeneous for estimating sample averages.

4.5. Confirmatory Factor Analysis

In order to check whether we are allowed to compare correlations and average scores in both countries, we used multigroup confirmatory factor analysis (MGCFA) with seven factors. For this analysis, we used the program Mplus 7.4 (Muthen & Muthen, 1998-2015). The usual χ^2 -based test for model fit is substantially affected by sample size [43]. Because we have large samples of observations, we use the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI). Both CFI- and TLI-indices are less vulnerable to sample size. Furthermore, we consider the Root Mean Square Error of Approximation (RMSEA) to assess model fit. The norms for acceptable fit are CFI and TLI > .90 and RMSEA < .08 [44-50].

Table 3 shows that both the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI) for Configural equivalence are above the norm of .90 and the Root Mean Square Error of Approximation (RMSEA) is below the norm of .08, which is an indication that in both countries the same 7 factor structure is found for the 35 items. The three indices are also in agreement with the norms for metric equivalence. This means that the factor loadings are equal in both countries, which is necessary for comparing correlations. The CFI, TLI and RMSEA are also in agreement with norms for scalar equivalence. This means that not only the factor loadings are equal, but also the intercepts. This allows us to compare the average scores of both countries, which is the aim of this study.

Norms for acceptable fit	RMSEA	CFI	TLI	
	<.08	>.90	>.90	
Configural equivalence: (the same 7 factor structure for the 35 items)	.062	.949	.943	
Metric equivalence: (equal factor loadings; necessary for comparing correlations)	.062	.948	.944	
Scalar equivalence: (equal factor loadings and intercepts; necessary for comparing averages)	.068	.932	.931	

Table 3: MGCFA 7 ICALT Scales South Korea (254) Netherlands (384)

So, we are allowed to compare the correlations and average scores in both countries.

4.6. Predictive Validity

We studied the predictive validity of the six ICALT-scales by computing the correlations between these scales and the observed engagement of the students. The correlations are presented in table 4.

basic skills			advanced skill	all 32 ICALT			
e l		activating differentiation students		teaching learning strategies	items		
South Korea	.61	.73	.69	.73	.49	.60	.75
Netherlands	.53	.67	.61	.62	.36	.39	.67

Table 4: Correlations with Student Engagement

According to Cohen correlations of .10 are small, and correlations of respectively .30 and .50, are medium and large [1]. The correlations between the quality of teaching and student engagement are large for the Korean teachers, and medium to large for Dutch teachers. So in both countries, better teaching skills go along with better student engagement. This is evidence for the predictive validity of the ICALT-scales.

5. Results

Scores on the ICALT-observations-scales may vary between 1 and 4. Scores ≤ 2 mean that the observed skill is not sufficient; scores between 2 and 3 indicate sufficient skill and scores ≥ 3 indicate that the skill is good. Table 5 presents the average scores of the South Korean and Dutch teachers in elementary education.

teaching skill	Korean teachers			Dutch to	Dutch teachers			T-value	sign.
	n	μ	sd	n	μ	sd			
safe learning climate	251	3.50	.57	384	3.60	.46	.20	2.43	.02
classroom management	247	3.44	.53	384	3.38	.50	12	1.44	.15
clear instruction	249	3.22	.63	384	3.27	.52	09	1.09	.26
activating students	237	3.11	.60	384	2.99	.56	21	2.52	.01
differentiation	251	2.63	.74	384	2.78	.72	.21	2.54	.01
teaching learning strategies	250	2.73	.78	384	2.23	.73	67	8.20	.0001
all 32 ICALT items	230	3.03	.51	384	2.92	.44	17	2.80	.005
student engagement	225	3.27	.66	384	3.11	.56	27	3.18	.001

Table 5: Average Scores of South Korean and Dutch Teachers in Elementary Education

Dutch teachers show on average good skills on basic skills like creating a safe and stimulating climate, efficient classroom management, and clear and structured instruction. On the advanced skills: activating students, and differentiation the Dutch teachers score, on average ample sufficient (≥ 2.5). On teaching learning strategies, the Dutch teachers score on average below ample sufficient (2.5). Dutch students show on average good involvement in the lessons.

South Korean teachers show on average good skills on creating a safe and stimulating climate, efficient classroom management, clear and structured instruction and intensive and activating teaching. The scores of Korean teachers on differentiation and teaching learning strategies are on average a bit better than ample sufficient (≥ 2.5). Korean students show on average good involvement in the lessons.

South Korean teachers do not differ significantly from Dutch teachers on: 'classroom management' and 'clear instruction'. Dutch teachers are little bit, but significantly better than South

Korean teachers on 'safe climate' (e.s.:.20) and 'differentiation' (e.s.:.21). South Korean teachers are a little bit better than Dutch teachers on 'activation students' (e.s.:.21) and clearly better on 'teaching learning strategies' (.67). 'Engagement of students' is a little bit better in South Korea than in the Netherlands (e.s.:.27). The differences between Dutch and South Korean teachers on 'safe climate' and differentiation are according to the criteria of Cohen and Lipsey are small [1, 2]. According the norms of Cohen and Lipsey the differences between activating students and teaching learning strategies are respectively small and medium size [51-55].

6. Discussion6.1. Sample Size

In this study we found small to medium size differences (Cohen's δ .35-50) in three teaching skills in favour of the South Korean teachers. Our sample was large enough to be able to see that these differences are significant. For finding effect sizes of >.15 significant on .05-level with a power of .80 two samples of each more than 550 teachers is needed. We should keep this in mind

for next studies into international comparisons.

6.2. Practical Implications

We know from OECD studies that the average scores of Dutch 15 year old students are almost a quarter of a standard deviation lower than the average scores of South Korean students.

Of course, we know that student's achievements are mostly caused by the students' intelligence and socio-economic background, but there is a lot of evidence that student's achievements are for 15-25% the result of the work of teachers. We also know from several studies that a growth in in the quality of teaching of about half a standard deviation goes along with an improvement of about 20% of a standard deviation in student results.

In this study we have concluded that South Korean teachers in elementary education are better than Dutch teachers on 'classroom management', 'intensive and activating teaching' and 'teaching learning strategies'.

Suppose we were able to coach Dutch teachers in such a way that their skills as a teacher improve, would the performance of their pupils come closer to the Korean pupils? This seems to be an interesting question for a follow-up study.

Acknowledgements

This study is partial output of two projects:

- 'Improving Teaching Skill by Classroom Observation and Analysis', which was supported by KRF research grant (2017S1A5A2A03067650)
- 'Come and Look in my Classroom', which was supported with a grant from the Dutch Labor Market Platform Primary Education

References

- Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- 2. Lipsey, M. W. (1990). Design sensitivity: Statistical power for experimental research (Vol. 19). sage.
- Van de Grift, W. J., Chun, S., Maulana, R., Lee, O., & Helms-Lorenz, M. (2017). Measuring teaching quality and student engagement in South Korea and The Netherlands. School Effectiveness and School Improvement, 28(3), 337-349
- Coleman, J. S. (1966). Equality of educational opportunity [summary report (Vol. 1). US Department of Health, Education, and Welfare, Office of Education.
- Hattie, J. (2008). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. routledge.
- 6. Hattie, J. (2012). Visible learning for teachers: Maximizing impact on learning. Routledge.
- Marzano, R. J. (2003). What works in schools: Translating research into action. ASCD.
- 8. Muijs, D., & Reynolds, D. (2017). Effective teaching: Evidence and practice. Third Edition, Sage.
- 9. Scheerens, J. (1992). Effective schooling: research theory and practice.
- 10. Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers

- and student achievement in the Chicago public high schools. Journal of labor Economics, 25(1), 95-135.
- 11. Bosker, R., & Witziers, B. (1996). The true size of school effects. In AERA Annual Meeting 1996: Research for Education in a Democratic Society.
- 12. Brandsma, H. P., & Knuver, J. W. M. (1989). Effects of school and classroom characteristics on pupil progress in language and arithmetic. International Journal of Educational Research, 13(7), 777-788.
- 13. Creemers, B. P. M. (1994). The Effective Classroom. London: Cassell. Data Analysis Methods.
- Cotton, K. (1995). Effective schooling practices: A research synthesis, 1995 update. Northwest Regional Education Laboratory.
- 15. Ellis, E. S., & Worthington, L. A. (1994). Research Synthesis on Effective Teaching Principles and the Design of Quality Tools for Educators. Technical Report No. 5.
- 16. Hanushek, E. A. (2011). The economic value of higher teacher quality. Economics of Education review, 30(3), 466-479.
- 17. Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. American economic review, 100(2), 267-271.
- 18. Houtveen, A. A. M., & Van de Grift, W. J. C. M. (2007). Effects of metacognitive strategy instruction and instruction time on reading comprehension. School effectiveness and school improvement, 18(2), 173-190.
- 19. Houtveen, T., & van de Grift, W. (2007). Reading instruction for struggling learners. Journal of Education for Students Placed at Risk (JESPAR), 12(4), 405-424.
- 20. Houtveen, A. A., van de Grift, W. J., & Brokamp, S. K. (2014). Fluent reading in special primary education. School Effectiveness and School Improvement, 25(4), 555-569.
- Houtveen, A. A. M., Van de Grift, W. J. C. M., & Creemers, B. P. M. (2004). Effective school improvement in mathematics. School Effectiveness and School Improvement, 15(3-4), 337-376.
- 22. Levine, D. U., & Lezotte, L. W. (1990). Unusually effective schools: A review and analysis of research and practice.
- 23. Levine, D. U., & Lezotte, L. W. (1995). Effective schools research. In J. A. Banks & C. A. M. Banks (Eds). Handbook of research on multicultural Education (pp. 525-547). New York, NY: Macmillan.
- 24. Marzano, R. J., Pickering, D., & Pollock, J. E. (2001). Classroom instruction that works: Research-based strategies for increasing student achievement. ASCD.
- 25. Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. The elementary school journal, 83(4), 427-452.
- 26. Purkey, S. C., & Smith, M. S. (1985). School reform: The district policy implications of the effective schools literature. The elementary school journal, 85(3), 353-389.
- Sammons, P. (1995). Key characteristics of effective schools: A review of school effectiveness research. B & MBC Distribution Services, 9 Headlands Business Park, Ringwood, Hants BH24 3PB, England, United Kingdom.
- 28. Van de Grift, W. (1985). Onderwijsleerklimaat en leerlingprestaties [Educational climate and school achievement]. Pedagogische Studiën, 62(10), 401–414.
- 29. van de Grift, W. (1990). Het onderzoek naar effectieve

- scholen. Pedagogische Studiën, 67(10).
- 30. Walberg, H. J., & Haertel, G. D. (1992). Educational psychology's first century. Journal of Educational Psychology, 84(1), 6-19.
- 31. Wijnstra, J., Ouwens, M., & Béguin, A. (2003). De toegevoegde waarde van de basisschool. Arnhem: CITOgroep.
- 32. Sanders, W. L., Wright, S. P., & Horn, S. P. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. Journal of personnel evaluation in education, 11, 57-67.
- 33. Houtveen, T., & van de Grift, W. (2011). Improving reading achievements of struggling learners. School Effectiveness and School Improvement, 23(1), 71-93.
- 34. Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. Review of educational research, 88(4), 547-588.
- 35. Van de Grift, W. J. C. M. (1994). Technisch rapport van het onderzoek onder 386 basisscholen ten behoeve van de evaluatie van het basisonderwijs [Report on the evaluation of 386 schools for primary education]. Inspectie van het Onderwijs, De Meern.
- Van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. Educational research, 49(2), 127-152.
- 37. Basisonderwijs, C. E. (1994). Zicht op kwaliteit: Evaluatie van het basisonderwijs: Eindrapport. De Meern: Inspectie van het basisonderwijs.
- 38. van de Grift, W., & Lam, J. (1998). Het didactisch handelen in het basisonderwijs. Tijdschrift voor onderwijsresearch, 23(2), 224-241.
- 39. Van de Grift, W. J. C. M., Van der Wal, M., & Torenbeek, M. (2011). Ontwikkeling in de pedagogische didactische vaardigheid van leraren in het basisonderwijs. Pedagogische Studiën, 88(6), 416-432.
- 40. van de Grift, W. J. (2014). Measuring teaching quality in several European countries. School effectiveness and school improvement, 25(3), 295-311.
- 41. van de Grift, W. J., Houtveen, T. A., van den Hurk, H. T., & Terpstra, O. (2019). Measuring teaching skills in elementary education using the Rasch model. School Effectiveness and School Improvement, 30(4), 455-486.

- 42. Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. psychometrika, 16(3), 297-334.
- 43. Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. Psychological bulletin, 103(3), 391-410.
- 44. Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural equation modeling: a multidisciplinary journal, 14(3), 464-504.
- 45. Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. Sociological methods & research, 36(4), 462-494.
- 46. Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural equation modeling, 9(2), 233-255.
- 47. Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural equation modeling: a multidisciplinary journal, 6(1), 1-55.
- 48. 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(3), 320-341.
- 49. Kline, R. B. (2005). Principles and Practice of Structural Equation Modeling (2nd Ed.). New York: Guilford.
- 50. Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. Psychometrika, 38(1), 1-10.
- 51. Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and psychological measurement, 20(1), 37-46.
- 52. Organisation for Economic Co-operation and Development. (2013). PISA 2012 results in focus: What 15-year-olds know and what they can do with what they know.
- OECD. (2014). PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science. (Volume I, Revised edition, February 2014), PISA, OECD Publishing.
- 54. OECD. (2016). PISA 2015 Results (Volume I): Excellence and Equity in Education. PISA, OECD Publishing.
- 55. OECD. (2019). PISA 2018 Results (Volume I): What Students Know and Can Do. PISA, OECD Publishing.

Copyright: ©2023 Wim van de Grift, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.