Abstract Title Page

Title: Effects of an Intensive Data-Based Decision Making Intervention on Teacher Efficacy

Authors and Affiliations:
Emmelien A. van der Scheer, Unversity of Twente
Adrie J. Visscher, University of Twente

Abstract Body

Background / Context:
Despite the emphasis on data-based decision making (DBDM) in educational policy in several countries, evidence regarding the intended effect of improved student achievement is still scarce (Campbell & Levin, 2009). Professional development programs (PDPs), designed to support schools and teachers with respect to the analysis and interpretation of data, infrequently lead to the desired effects (Carlson, Borman, & Robinson, 2011). The Dutch Inspectorate of Education (2014) reported that although Dutch teachers have improved their ability to analyze data, they are yet to adapt their instruction towards the needs of students, as shown by the data analyzed. Possibly, teachers need more support, beyond a training course in data analysis skills, to adapt instruction. However, such skills are advanced teaching skills not mastered well by a considerable proportion of teachers (Van de Grift, 2007).

A teacher's efficacy (TE) reflects whether a teacher thinks he is able to support student learning, therefore, a high sense of efficacy might be important for working in a data-based way (Bruce, Esmonde, Ross, Dookie, & Beatty, 2010). Teachers with a high sense of efficacy are confident about their ability to enhance student learning. Teachers with a low sense of efficacy predominantly attribute student learning to other factors than themselves (Bruce et al., 2010). Moreover, teachers with a higher sense of efficacy are more likely to implement new teaching practices and when doing so will persevere, if confronted with difficulties (Wolters & Daugherty, 2007). Therefore, a higher sense of teacher efficacy might promote the implementation of new (DBDM) practices in the classroom (Stein & Wang, 1988).
TE is mainly formed on the basis of a teacher's own experiences within the classroom (mastery experiences), observing well-performing peer teachers who succeed (vicarious experiences), and verbal persuasion by significant others (e.g. school leaders) (Bandura, 1997).

Only a small number of teacher efficacy studies was conducted with either an experimental, or a longitudinal research design (Klassen, Tze, Betts, & Gordon, 2011). As a result, little is known about the extent to which teacher efficacy can be improved through interventions (Henson, 2001; Klassen et al., 2011).

Purpose:
The purpose of this study is to examine the effects of a DBDM intervention on teacher efficacy. This study incorporates an experimental longitudinal research design, as teacher efficacy was measured at three stages: prior to the intervention, immediately after it, and a school year later. The main question answered in this study is:
What is the effect of an intensive DBDM intervention on teachers' efficacy?

Setting:
This study was conducted with grade 4 teachers from regular primary schools in the Netherlands.

Participants:
Grade 4 teachers from 60 primary schools in the Netherlands agreed to participate, and were randomly allocated (at the school level) to either treatment group 1 (30 classes), or treatment group 2 (30 classes). Teachers in treatment group 1 (39 teachers from 30 schools) participated in the DBDM-intervention during the school year 2013-2014. Teachers allocated to treatment group 2 (36 teachers from 30 schools) were used as control group teachers during the school year 2013-2014, and were exposed to the intervention during the school year 2014-2015.

Participating teachers taught a grade 4 class (9-10 year old students), or a multi-grade classroom including both grade 4 students and students from other grades.
A number of teachers in both treatment groups dropped out for various reasons (e.g. maternity leave, long term illness, organizational reasons). The remaining number of participating teachers are presented in Table 1.

(Please insert Table 1 here)

Intervention:
Teachers were exposed to a DBDM intervention, which included seven meetings (a total of 36 hours of contact time), and four coaching sessions during which they were coached in the classroom on how to implement DBDM. The intervention was designed on the basis of five criteria considered important for effective professional development by Desimone (2009): active learning, a considerable duration of the intervention, content focus, coherence between the intervention and (school)policy, and collective participation. During the seven meetings, teachers were mainly trained in data-use, in how to formulate performance and learning goals, and in how to draw up an instructional plan. The implementation of DBDM in the classroom was strongly emphasized in meeting four and sixth, and in four individual classroom coaching sessions.

During these meetings, teachers watched a videotape of their own lesson (reduced to approximately 20 minutes of the original lesson). In the fourth meeting this was done in a teacher pair (two teachers from different schools), whereas a videotaped lesson of each teacher was shown to all teachers who attended meeting six. After the lesson had been shown, the teachers discussed the various lesson phases.

During four individual coaching sessions each teacher received feedback from the trainer on how they implemented DBDM in his/her classroom. After the observed lesson, the trainer and teacher discussed the various lesson phases. Teachers first presented their opinions regarding their strengths and weaknesses for each lesson phase. Next, the trainer provided his own opinion on each of the phases. Finally, the teacher and trainer discussed how the teacher could improve his/her DBDM-implementation.

Especially during these meetings and coaching sessions the teacher efficacy may have been impacted. Teachers could have experienced *mastery experiences,* as explicit feedback was provided on a person's performance, and reflection on his functioning was stimulated. Teachers also observed multiple peer teachers by means of video fragments, in which *vicarious experiences* could have taken place. Furthermore, the trainer may have persuaded *verbally* teachers when they received feedback during the individual coaching sessions. An overview of the intervention is provided in Table 2.

(Please insert Table 2 here)

Research Design:
Teachers completed an online questionnaire prior to (T1), after the first intervention year (T2), and after the second intervention year (T3). This questionnaire included, among others, questions on teacher background characteristics, and a Dutch translation of the long form of the Teachers' Sense of Efficacy Scale (TSES) (Tschannen-Moran & Woolfolk-Hoy, 2001). The TSES measures teacher efficacy with three constructs: teacher efficacy with respect to classroom management, instructional strategies, and student engagement (Tschannen-Moran & Woolfolk-Hoy, 2001). We changed the response categories from a 9-point Likert- scale to a 5-point Likert-scale as we considered a 5-point scale more feasible for our respondents.

Data Collection and Analysis:
A mean score per teacher, per scale, per measurement moment, was calculated to determine a teacher's efficacy on each of the scales (Tschannen-Moran & Woolfolk-Hoy, 2001). To evaluate the effect of the intervention on teachers' efficacy scores during the *school year 2013-2014* a multivariate analysis of variance (MANOVA), and a multivariate analysis of covariance (MANCOVA) were performed with SPSS (version 22). The assumptions for a MAN(C)OVA were not met for the classroom management scale, the classroom management scale was therefore analyzed separately using a nonparametric Mann-Whitney test. The MAN(C)OVA was performed with the instructional strategies and student engagement scores prior to the intervention (T1) as covariates, participation in the intervention as a factor, and the instructional strategies and student engagement scores after the intervention as dependent variables (T2).

For the analysis of the effect of the intervention in the *school year 2014-2015* on treatment group 2 teachers, their T3 scores were compared with their T2 scores, by means of paired samples t-tests for the instructional strategies scale, and the student engagement scale. We performed a Wilcoxon signed-rank test on the classroom management scale. The same tests were executed to investigate whether the intervention effect in treatment group 1 remained the same during the school year 2014-2015.

Results:
Intervention year 2013 - 2014
Teachers in treatment group 1 received the intervention in the intervention year 2013-2014. Table 3 presents the mean scores per scale, per treatment group prior to (T1), and after (T2) the intervention year, and the differences between T1 and T2. A graphical representation of these results is presented in Figure 1.

(Please insert Table 3 here)
(Please insert Figure 1 here)

As shown in Table 3 and Figure 1, for both the instructional strategies scale and the student engagement scale, teachers in treatment group 1 reported after the intervention a higher sense of efficacy compared to teachers in treatment group 2. The multivariate (MANOVA) result was statistically significant for participation in the intervention (Pillai's Trace = .14, $F(2, 59) = 4.78$, $p < .05$), indicating that teachers who had been exposed to the intervention had a significantly higher sense of efficacy after the intervention. This was found for both the instructional strategies scale ($F(1, 60) = 8.18$, $p < .05$), and the student engagement scale ($F(1, 60) = 8.23$, $p < .05$). With the inclusion of the T1 scores as covariates, the MANCOVA showed that both the multivariate effect (Pillai's Trace = .34, $F(2, 57) = 14.81$, $p < .05$), and the univariate effects were statistically significant (for instructional strategies, $F(1, 58) = 26.57$, $p < .05$, and for student engagement, $F(1, 58) = 17.90$, $p < .05$). No significant effect of the intervention was found for the classroom management scale ($U = 369.00$, $z = -1, 58$, $p > .05$).

Intervention year 2014 - 2015
Teachers in treatment group 2 received the intervention during the intervention year 2014-2015. In Table 4, we present the mean scores across scales and treatment groups at T1, T2 and T3 as well as the scores for the difference between T2 and T3. Moreover, the Table presents the statistically significant differences between both measurement moments. The graphical representation of these results can be found in Figure 2.

(Please insert Table 4 here)
(Please insert Figure 2 here)

Table 4 and Figure 2 show that for all the scales, treatment group 2 teachers reported a higher sense of teacher efficacy at T3 than at T2 (a difference ranging from .08 to .28). By means of paired samples t-tests, it was found that the gain in teacher efficacy between T2 and T3 for teachers in treatment group 2 was statistically significant for instructional strategies ($t(14) = 2.74$, $p < .05$), and for student engagement ($t(14) = 5.48$, $p < .05$). However, the mean rank for classroom management ($z = -1.10$, $p > .05$) did not differ significantly between T2 and T3.

As shown in Table 4, teachers in treatment group 1 had a slightly lower sense of efficacy at T3 in comparison to T2 for all scales. However, paired sample t-tests showed that the scores at both points in time were comparable as no significant differences were found for the instructional strategies scale ($t(20) = -.58$, $p > .05$), and the student engagement scale ($t(20) = -.36$, $p > .05$). The Wilcoxon signed-rank test also showed for the classroom management scale ($z = -.68$, $p > .05$), that the scores at T2 and T3 were comparable for treatment group 1.

Conclusions:

This study, based on a strong research design, shows that a teacher's sense of efficacy regarding instructional strategies and student engagement can be improved significantly by means of an intensive intervention, and that this effect persists at least a year after the intervention was concluded. This is the first TE study that incorporated a long randomized controlled trial in which such positive intervention effects on teacher efficacy were found. These effects were not found for the classroom management scale. The fact that Dutch in-service primary school teachers after the intervention felt more confident about their instructional strategies, and their ability to engage students is a significant finding as teacher efficacy is generally considered to have become stable after several years of teaching (Tschannen-Moran & Woolfolk-Hoy, 2001), therefore hard to change, and, if changed, these changes are assumed not to last (Klassen et al., 2011).

Desimone (2009) emphasizes that the change of teacher beliefs and attitudes is the first step in changing teacher practice and student achievement. Teacher efficacy can play an important role in the implementation of an intervention that incorporates complex teaching skills (like DBDM). Teachers with a higher sense of efficacy have more confidence in their ability to influence student performance. Moreover, they are more willing to try new teaching practices, to set more challenging goals, to put more effort in teaching, and they are also more resilient (Tschannen-Moran et al., 1998). For the implementation of DBDM, which requires complex teaching skills (like instructional differentiation based on student progress data), a high sense of teacher efficacy may be beneficial.

TE reflects how a teacher *perceives* his/her professional competences, which is not necessarily the same as his objective professional competence in the classroom. Research is needed on the assumed beneficial effects of improved levels of teacher efficacy as such studies are not available yet. In future studies, we should clarify whether improved TE goes together with improved teacher performance in the classroom, and with improved student performance. Finally, research in other teacher contexts is needed to clarify whether this (kind of) intervention has similar effects on teachers in other grades, and other school types (e.g. on secondary school teachers).

Appendices
Appendix A. References

Bandura, A. (1997). Self-efficacy: the exercise of control. New York: W.H. Freeman and Company.

Bruce, C. D., Esmonde, I., Ross, J., Dookie, L., & Beatty, R. (2010). The effects of sustained classroom-embedded teacher professional learning on teacher efficacy and related student achievement. Teaching and Teacher Education, 26, 1598-1608.

Campbell, C., & Levin, B. (2009). Using data to support educational improvement. Educational Assessment, Evaluation and Accountability, 21, 47-65.

Carlson, D., Borman, G. D., & Robinson, M. (2011). A Multistate District-Level Cluster Randomized Trial of the Impact of Data-Driven Reform on Reading and Mathematics Achievement. Educational Evaluation and Policy Analysis, 33(3), 378-398. doi: 10.3102/0162373711412765

Desimone, L. M. (2009). Improving impact studies of teacher's professional development: toward better conceptualizations and measures. Educational Researcher, 38(3), 181-199. doi: 10.3102/0013189X08331140

Henson, R. K. (2001). The effects of participation in teacher research on teacher efficacy. Teaching and Teacher Education, 17, 819-836. doi: 10.1016/S0742-051X(01)00033-6

Inspectorate of Education. (2014). De Staat van het Onderwijs. Onderwijsverslag 2012/2013. Utrecht: Inspectorate of Education.

Klassen, R. M., Tze, M. C. V., Betts, S. M., & Gordon, K. A. (2011). Teacher Efficacy Research 1998–2009: Signs of Progress or Unfulfilled Promise? Educational Psychology Review, 23, 21-43.

Stein, M. K., & Wang, M. C. (1988). Teacher Development and School Improvement: The Process of Teacher Change. Teaching and Teacher Education, 4(2), 171-187.

Tschannen-Moran, M., & Woolfolk-Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. Teaching and Teacher Education, 17, 783-805.

Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher Efficacy: Its Meaning and Measure. Review of Educational Research, 68(2), 202-248. doi: 10.3102/00346543068002202

Van de Grift, W. J. C. M. (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.

Wolters, C. A., & Daugherty, S. G. (2007). Goal Structures and Teachers' Sense of Efficacy: Their Relation and Association to Teaching Experience and Academic Level. Journal of Educational Psychology, 99(1), 181-193. doi: 10.1037/0022-0663.99.1.181

Appendix B. Tables and Figures

Table 1
The number of teachers and classes in each intervention year, per treatment group and measurement moment

Group	Year	T1		T2		T3	
		N classes	N teachers	N classes	N teachers	N classes	N teachers
Treatment group 1	2013/2014	30	39	25	32		
	2014/2015			25	32	19	21
Treatment group 2	2013/2014	30	36	27	30		
	2014/2015			20	23	15	15

Table 2
Intervention content

Month	**Duration**	**Meeting**	**Content**
September	8 hours	Meeting 1	- The background and meaning of DBDM - Explanation CITO-assessments - Explanation use and interpretation SMS - Analysis and interpretation own data with a SMS by means of a protocol - Individualized feedback on the SMS-protocol
September	8 hours	Meeting 2	- Drawing up an instructional plan - Formulating SMART goals - Individualized feedback on instructional plan
October	4 hours	Meeting 3	- Explanation Dutch core goals - The mathematics learning progression - Instructional models - Drawing up personal improvement goals
Coaching session 1			
November	4 hours	Meeting 4	- Observation mastery lessons - Observation videotaped lessons in teacher pairs - Peer teachers provide feedback on videotaped lessons - Feedback on how to differentiate instruction in the classroom
Coaching session 2			
February	4 hours	Meeting 5	- As in meeting 1
Coaching session 3			
April	4 hours	Meeting 6	- As in meeting 4
Coaching session 4			
June	4 hours	Meeting 7	- As in meeting 1

Table 3

Mean scores per treatment group for T1 and T2, and the difference between T1 and T2

			T1		T2		T2-T1	
		N	M	SD	M	SD	M	SD
Classroom management	Treatment group 1	32	3.95	.49	4.21	.47	.26	.40
	Treatment group 2	30	4.22	.44	4.31	.50	.10	.42
Instructional strategies	Treatment group 1	32	3.62	.42	4.13	.42	.52	.35
	Treatment group 2	30	3.76	.54	3.83	.42	.07	.37
Student engagement	Treatment group 1	32	3.46	.36	3.88	.38	.42	.33
	Treatment group 2	30	3.52	.49	3.60	.37	.08	.34

** $p < .05$

Table 4

Mean scores per treatment group for T1, T2 and T3, and the difference between T2 and T3

			T1		T2		T3		T3-T2	
		N	M	SD	M	SD	M	SD	M	SD
Classroom	Treatment group 1	21	3.96	.51	4.26	.47	4.20	.61	-.06	.52
management	Treatment group 2	15	3.97*	.37	4.16	.46	4.23	.70	.08	.40
Instructional	Treatment group 1	21	3.63	.43	4.17	.46	4.12	.46	-.05	.37
strategies	Treatment group 2	15	3.60*	.43	3.75	.31	3.95	.36	.20**	.28
Student	Treatment group 1	21	3.49	.39	3.90	.43	3.88	.37	-.03	.35
engagement	Treatment group 2	15	3.34*	.39	3.44	.20	3.72	.24	.28**	.20

* The T1 scores of treatment group 2 were calculated based on the scores of 13 teachers (the scores of two teachers were not available).

** $p < .05$

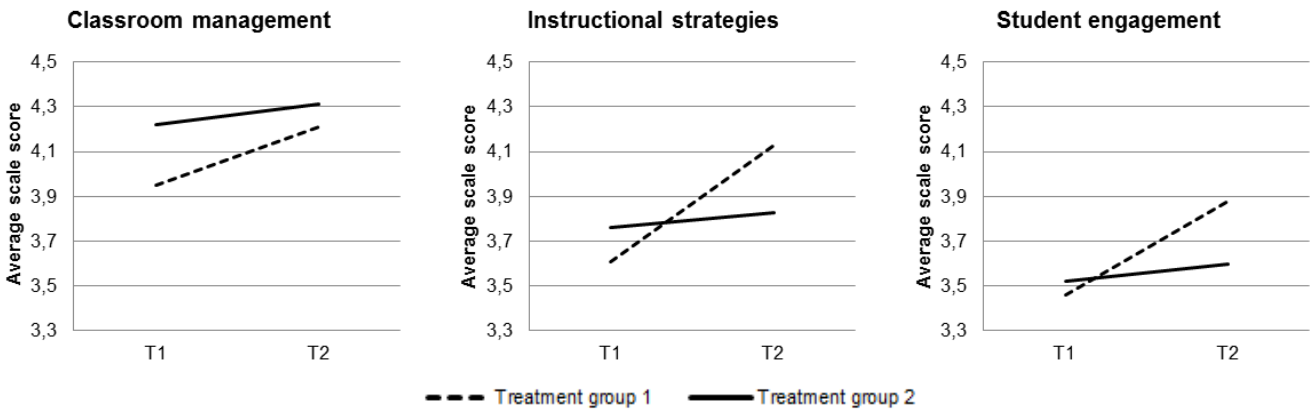

Figure 1. Mean scores per treatment group for T1 and T2 in school year 2013-2014.

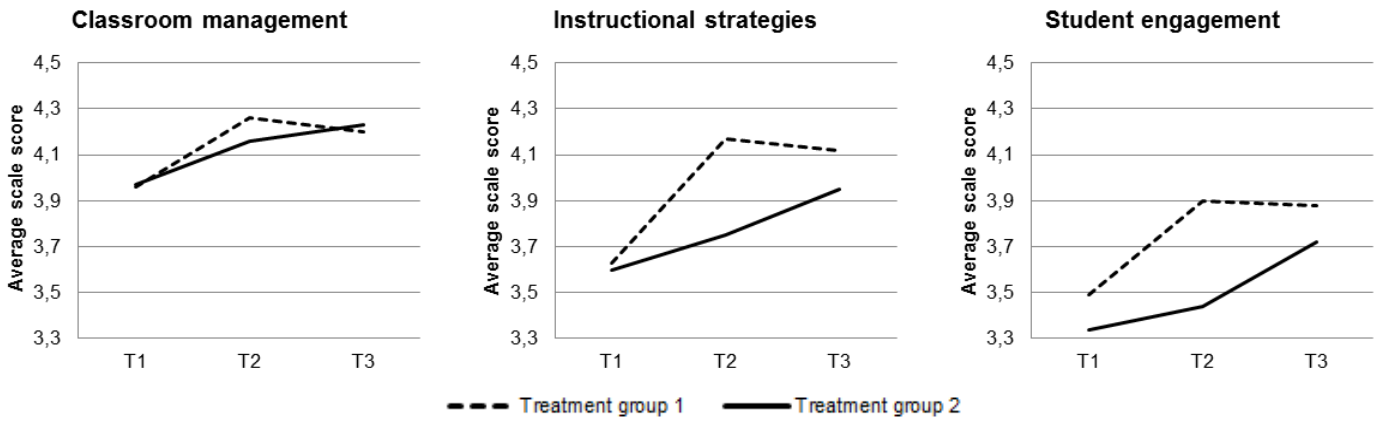

Figure 2. The mean scores per treatment group for T1, T2 and T3.