

Before, During, and After Examination: Development of Prospective Preschool Teachers' Mathematics-Related Enjoyment and Self-Efficacy

Sigrid Blömeke^a, Oliver Thiel^b and Lars Jenßen^c

^aCentre for Educational Measurement, University of Oslo (CEMO), Oslo, Norway; ^bSeksjon for Matematikk, Dronning Mauds Minne Høgskole for Barnehagelærerutdanning, Trondheim, Norway; ^cDepartment of Education and Psychology, Freie Universität Berlin, Berlin, Germany

ABSTRACT

This article examines the stability of Norwegian prospective preschool teachers' enjoyment of mathematics and their mathematics-related self-efficacy before, during, and after a teacher-education examination. In addition, the stability of the two constructs across countries was examined through a comparison with Germany. The data revealed partial stability (technically speaking, metric invariance) of enjoyment but not of self-efficacy. Self-efficacy increased significantly before and after the examination without decreasing enjoyment, which may be a result of increased learning time. Prior mathematical knowledge predicted the level and development of enjoyment and self-efficacy in both countries. Many Norwegian students reported low mathematics-related enjoyment and self-efficacy, including negative developments. It may be important to provide positive experiences of mathematical activities during preschool teacher education.

ARTICLE HISTORY

Received 12 December 2016
Accepted 29 August 2017


KEYWORDS

Early childhood education; preschool teacher; mathematics; longitudinal study; attitudes

Introduction

Fostering early mathematical literacy has come to attention in Norway in the most recent national framework for preschool education set by the government (Mosvold & Sæbbe, 2015) because research has revealed that early mathematical literacy is decisive for later school achievement in mathematics (Duncan et al., 2007; Krajewski & Schneider, 2009). In many countries, preschool teachers are therefore nowadays expected to ensure that children experience numbers, shapes, and patterns. However, the state of research suggests that attitudinal antecedents of such activities may be lacking. Preschool teachers seem to have difficulty recognizing the value of mathematics and believe more strongly in their responsibility to foster the socio-emotional development of children (Kowalski, Pretti-Frontczak, & Johnson, 2001; Lee & Ginsburg, 2007; Tietze, Roßbach, & Grenner, 2005; Varol, 2013). Furthermore, they judge their own efficacy in implementing mathematics-related activities in preschool as rather low (Ginsburg, Lee, & Boyd, 2008).

Such low domain-specific enjoyment and self-efficacy are unfortunate because studies reveal not only that these characteristics are related to the quantity and quality of preschool teachers' mathematics-related interaction with children but also that teachers with stronger domain-specific enjoyment and self-efficacy are more *actively* interested in creating corresponding learning opportunities

CONTACT Sigrid Blömeke  Sigridbl@cemo.uio.no  Centre for Educational Measurement, University of Oslo (CEMO), Oslo, Norway

for children (Guo, Piasta, Justice, & Kaderavek, 2010; Richardson, 1996; Singh, Granville, & Dika, 2002). Such an active interest would be highly relevant in the Norwegian preschool system, which follows a more “social pedagogy” than a “pre-primary tradition” (OECD, 2006).

Play represents the norm in Norway but teachers are nevertheless required to support children’s enjoyment of mathematics (*matematikkglede*) and to combine play with learning (Alvestad & Samuelsson, 1999). In contrast to the structured pre-primary model often implemented in the US, the challenge for Norwegian preschool teachers is thus to use daily activities for fostering early mathematics informally and through care (Nakken & Thiel, 2014). Mosvold and Sæbbe (2015) provide a description of what it means to “teach” mathematics in a Norwegian preschool and the challenges connected to that. Attitudes may play an even stronger role in such a context.

Against this background, the purpose of the study is to examine prospective preschool teachers’ mathematics-related enjoyment (M_JOY) and their mathematics-related self-efficacy (M_SELF) in Norway. Examining these attitudinal antecedents of behavior may increase our understanding of preschool teachers’ interaction with children, in particular the absence or presence of mathematics-related activities. We are particularly interested in examining the *stability or variability* of prospective preschool teachers’ M_JOY and M_SELF *under different context conditions*, namely, over time and across countries, including *predictors* of favorable patterns and developments. Such information gives us deeper insight into the nature of M_JOY and M_SELF and may point to ways in which to address prospective preschool teachers’ training needs during teacher education.

In the following, as a first step, we will provide a short overview of the state of research on M_JOY and M_SELF, as well as on the stability or variability of these constructs, before we specify our research questions and hypotheses. In the next section, we describe the methodological approach used in this study, including study design, sample, measures, and data analysis. The results are presented and then discussed and conclusions are drawn in the final section.

Mathematics-related Enjoyment and Self-efficacy

Enjoyment and self-efficacy are antecedents for a related behavior—in this case, dealing with mathematics as a precondition for the implementation of mathematical activities in preschool (Bandura, 1977; Fishbein & Ajzen, 1975). *Mathematics-related self-efficacy* can be defined as “an individual’s confidence in her or his ability to successfully perform a particular math-related task or problem” (Hackett & Betz, 1989, p. 262). *Mathematics-related enjoyment* describes motive, that is, why somebody would engage in mathematical activities for their own sake and not as means to gain other rewards (Deci & Ryan, 1985).

Research suggests that domain-specific self-efficacy and enjoyment are closely related to corresponding knowledge (Hackett & Betz, 1989; McMullan, Jones, & Lea, 2012). Higher knowledge increases an individual’s confidence in his or her ability to do something and makes the task at hand more enjoyable. Thus, it should be possible to explain preschool teachers’ levels of M_JOY and M_SELF through their level of mathematical content knowledge (MCK). Several studies have confirmed, in fact, medium or strong correlations in these respects for a broad range of groups, including preschool teachers (Bates, Latham, & Kim, 2011; Tsamir, Tirosh, Levenson, Tabach, & Barkai, 2015). More studies have investigated M_SELF than M_JOY.

Larger studies specifically about Norwegian preschool teachers’ MCK level or mathematics-related attitudes are lacking (Alvestad, Johansson, Moser, & Søbstad, 2009; Sommersel, Vestergaard, & Larsen, 2013). However, it might not be far-fetched to assume that both characteristics are not optimally developed given what else we know about the Norwegian preschool context. Norwegian preschool curricula have traditionally been—even in a Nordic context, for example in comparison with Sweden (Alvestad & Samuelsson, 1999)—quite reluctant to require the fostering of cognitive development by preschool teachers in general and even more to fostering the development of early mathematical literacy. The latter has only recently come to attention and the curricula have since been updated (Mosvold & Sæbbe, 2015). Still, opportunities to learn mathematics during

preschool teacher education are scarce, and Norwegian preschool teachers struggle with the implementation of mathematical activities in preschool (Mosvold, 2012; Mosvold, Bjuland, Fauskan-ger, & Jakobsen, 2011). We are not aware of studies that have examined the attitudinal antecedents of such activities in Norway.

Stability and Variability of Self-efficacy and Enjoyment

Cognitive characteristics such as knowledge and personality traits such as the so called “Big Five” are typically regarded as relatively stable over short periods of time and across various situations (Liebert & Liebert, 1998; McCrae & Costa, 1996). In contrast, behavior and affects are typically hypothesized to be less stable and to vary more strongly depending on the situation (Anastasi, 1983). Results regarding research into attitudes such as enjoyment or about self-efficacy are less clear. Whereas these characteristics often are assumed to be stable and can be changed only through strong cognitive dissonance, other researchers assume variability. In an extreme version (Wilson & Hodges, 1992), it is hypothesized that attitudes are always constructed in a situation, meaning that an underlying stable trait does not exist at all.

Technically speaking, stability of characteristics over different time points or across countries means that they have to have the same measurement characteristics each time and in each country. Only then can comparisons of their means or their relationships to other constructs make sense across the different contexts (Van de Vijver & Leung, 1997). The same dimensionality, factor loadings, and intercepts of a construct at different time points or in different countries would indicate stability (also called “measurement invariance”; see Horn & McArdle, 1992; Meredith, 1993). Very few studies have examined the stability or variability of (prospective) preschool teachers’ characteristics in these respects. A study in Germany revealed that prospective preschool teachers’ mathematics anxiety was stable over the course of three weeks if the context conditions were not varied (Jenßen, Dune-kacke, Eid, & Blömeke, 2015). How this would look like when the contexts differ is unknown.

With respect to the M_JOY and M_SELF of (prospective) preschool teachers, we are not aware of any research into their stability or variability. Research on self-reported measures in general (Laschke & Blömeke, 2016; Rutkowski & Svetina, 2014) suggests that it is often possible to confirm the same dimensionality (“configural invariance”) even if contexts are different. This means, in our case, that the same items would have to be associated with M_JOY and M_SELF at each time point or in each country, and such a result would mean that the internal structure of the constructs is the same.

Similarly, it is typically possible to confirm that the factor loadings of items used to assess a construct are invariant across different time points or countries (“metric invariance”). This means that the unit of measurement is identical, and it is possible to compare relationships to other constructs across different contexts.

In contrast, the general state of research on self-reported measures suggests that it is typically *not* possible to confirm scalar invariance across different contexts because the item intercepts are not invariant as a result of respondents’ different frames of reference and response biases or differences in understanding (Van de Vijver & Tanzer, 2004). One should be careful with comparisons of means in this case.

We varied the context conditions of time points by including two extreme situations: a regular learning situation during a preschool teacher education class on teaching mathematics with nothing at stake and a high-stakes situation one week later when the same students had to take their final examinations. It is known from other studies that the stressful experience of such a situation may reveal variability in personal characteristics—at least with respect to emotions (Spielberger & Reheiser, 2009). To what extent this also applies to M_JOY and M_SELF is an open question. For heuristic reasons, we added a third situation approximately 3–5 weeks after the examination when a small subsample of these students had received their grades.

The country context is varied by including Germany as well as Norway. Germany belongs to the same social pedagogy preschool tradition as Norway but trains preschool teachers mostly at

vocational schools following middle school (i.e., at the secondary or post-secondary school level; Statistisches Bundesamt, 2014), whereas preschool teacher education in Norway takes place at pedagogical colleges that are part of the higher education sector.

Research Questions and Hypotheses

- (1) The first research question addresses whether Norwegian prospective preschool teachers' M_JOY and M_SELF are stable across different time points; in particular, when they are about to take an examination compared to one week before and three to five weeks after this point in time. In line with the general state of research on measurement invariance, we hypothesize that configural longitudinal invariance in terms of dimensionality and metric longitudinal invariance in terms of equal factor loadings exist but not scalar invariance in terms of equal intercepts (H1).
- (2) The second research question addresses whether prospective preschool teachers' M_JOY and M_SELF are stable across Norway and Germany. We hypothesize, again, that configural and metric invariance exist but not scalar invariance (H2).
- (3) The third question addresses whether prior content knowledge in mathematics is able to predict prospective preschool teachers' M_JOY and M_SELF in Norway and Germany, as well as the development of these characteristics at different time points in Norway. We hypothesize that higher MCK is related to higher M_JOY and M_SELF in Norway and Germany (H3a) and that it is related to positive, or at least stable, non-decreasing developments of the two characteristics before and after the examination (H3b).
- (4) A final research question is exploratory in nature. Do all Norwegian prospective preschool teachers develop in the same way at different time points or do subgroups with different developmental processes exist? Such details would provide information about whether it is necessary to adapt teacher education to different needs.

Methodology

Samples

Participants in the longitudinal study were 225 full- and part-time student teachers from one preschool teacher education program in Norway. They were, on average, 27 years old (min = 20, max = 54 years), and 85% were female. Independent of their status as full- or part-time students, all had undergone teacher education equivalent to approximately two years of full-time study. Participants had, on average, a mathematics grade of 3.4 in their final year of schooling (min = 1.0 as the worst grade; max = 6.0 as the best grade). Approximately three-quarters had entered preschool teacher education after three years of high school, which is equivalent to general study competence.

At the first time point of the survey, one week before the semester's final examination, 197 students participated; 135 participants took part at the second time point, on the examination day itself. They filled in the survey right after they had taken their examination. Only 26 could be reached at the third time point about 3–5 weeks after the examination, during the summer break when students had received their grades. This survey was given online to the students whereas the first two time points used paper-and-pencil versions.

Despite the high drop-out rate, we can point out that a selection bias with respect to predicted prior knowledge in mathematics did not exist. We used the school grade in mathematics as a proxy, and its average was nearly identical for the groups surveyed at the three time points. Differences of 0.1 in students' grades were not significant. Nor did a significant gender difference exist between the three groups.

Participants in the cross-sectional study in Germany were 354 full-time prospective preschool teachers from five vocational schools located in several federal states. Participants were, on average, 23 years old (min = 17, max = 46), and 83% were female. About 42% of the sample were in the first year of teacher education, 33% in the second year, and 26% in the third and last year. Participants had, on average, a mathematics grade of 3.2 in their final year of schooling (min = 6.0 as the worst grade; max = 1.0 as the best grade—note that the scale is reversed compared to the Norwegian grading system!). A paper-and-pencil version of the survey was utilized.

Measures

In both countries, and at all time points, M_JOY was surveyed using five items (for the wording of these items, see Table 1), extending a set of items already applied in TEDS-M (Tatto et al., 2012) and in studies investigating preschool teacher education students in Germany (Blömeke, Dunekacke, & Jenßen, 2017). They were translated into Norwegian for the purposes of the present study. Participants had to rate these items on the original six-point Likert scale ranging from “strongly disagree” to “strongly agree.”

If one or two residual correlations were allowed between items that were more similar to each other than to other items as a result of wording, the scale fit well with the Norwegian data at the first two time points (first: $X^2_{(4)} = 6.3$, $p = .18$; CFI = .99, TLI = .98; RMSEA = .05, 90% CI [.00, .13], $p = .38$; SRMR = .03; second: $X^2_{(3)} = 4.5$, $p = .21$; CFI = 1.00, TLI = .98; RMSEA = .06, 90% CI [.00, .17], $p = .34$; SRMR = .02) and also at the third with a small sample size ($X^2_{(4)} = 4.8$, $p = .31$; CFI = .99, TLI = .97; RMSEA = .09, 90% CI [.00, .32], $p = .34$; SRMR = .03). The fit was good in Germany, too ($X^2_{(4)} = 2.5$, $p = .64$; CFI = 1.00, TLI = 1.01; RMSEA = .00, 90% CI [.00, .07], $p = .86$; SRMR = .01).

Table 1 reveals that the factor loadings were mostly of sufficient size—between 0.78 and 0.96—at the three time points in Norway, as well as at the one measurement point in Germany (between 0.74 and 0.92). However, in Norway the item “It is hard to enjoy mathematics” had a substantially lower loading at all three time points than the other four items.

M_SELF was surveyed with four items based on a well-established scale assessing teacher self-efficacy in general (Schwarzer & Jerusalem, 1995) but adjusted for the target population of the present study and translated into Norwegian. Participants had to rate the four items on the original four-point Likert scale ranging from “strongly disagree” to “strongly agree.”

The scale fit well with the data at all three time points in Norway (first: $X^2_{(2)} = 0.7$, $p = .69$; CFI = 1.00, TLI = 1.02; RMSEA = .00, 90% CI [.00, .11], $p = .80$; SRMR = .01; second: $X^2_{(2)} = 1.1$, $p = .56$; CFI = 1.00, TLI = 1.01; RMSEA = .00, 90% CI [.00, .15], $p = .66$; SRMR = .01; third: $X^2_{(2)} = 1.3$, $p = .53$; CFI = 1.00, TLI = 1.02; RMSEA = .00, 90% CI [.00, .34], $p = .56$; SRMR = .01). The fit of the scale to the data was good in Germany, too ($X^2_{(2)} = 2.6$, $p = .28$; CFI = 1.00, TLI = 1.00; RMSEA = .03, 90% CI [.00, .13], $p = .50$; SRMR = .01).

Table 2 documents the factor loadings in both countries. Throughout they were of sufficient size. Only the factor loading of the item “Although it is exhausting, I am able to push myself and to accomplish high achievements in mathematics” was slightly lower at the second time point in both Norway and Germany.

Table 1. Standardized factor loadings of the M_JOY items in Norway and Germany.

Enjoyment of mathematics (M_JOY)	Norway: First/second/third time point	Germany
5 Mathematics is boring (R)	.87/.81/.88	.77
2 Mathematics is enjoyable	.83/.96/.93	.92
1 Mathematics offers the possibility to enjoy discoveries	.83/.78/.86	.81
4 Mathematics leads to enjoyable experiences	.78/.79/.90	.74
3 It is hard to enjoy mathematics (R)	.67/.51/.47	.81

Note: Items sorted according to the size of the factor loadings at the first time point in Norway, R = reversed.

Table 2. Factor loadings of the M_SELF items at the three time-points.

Mathematics-related self-efficacy (M_SELF)	Norway: First/second/third time point	Germany
2 Although I have not solved mathematical problems for a long time, I am nevertheless able to solve such problems when I encounter them.	.78/.82/.91	.79
4 Although it is exhausting, I am able to push myself and to accomplish high achievements in mathematics.	.77/.67/.80	.68
1 I am sure that I can solve difficult mathematical problems as well.	.76/.83/.87	.82
3 I know different ways to solve difficult mathematical problems.	.74/.86/.94	.81

Note: Items sorted according to the size of the factor loadings at the first time point in Norway.

The final school grade in mathematics was used as a proxy for prospective preschool teachers' knowledge in mathematics, which they had brought with them into preschool teacher education. It was used as a within-country predictor of mathematics-related self-efficacy and enjoyment. Note that the means of the grades should not be compared across Germany and Norway due to differences in the respective national mathematics curricula.

Data Analysis

First, the fit of the scales to the data was examined for each time point and country using confirmatory factor analysis (CFA). Absolute and incremental fit indices were used to evaluate fit. The chi-square test indicates whether the covariance matrix implied by the scale model is equal to the population covariance matrix. The root square mean error of approximation (RSMEA) and the standardized root mean square residual (SRMR) estimate whether a model reproduces well the observed covariance or correlation matrix (estimates of $< .08$ point to a good fit and $< .05$ to a very good fit; Hu & Bentler, 1999). The Comparative Fit Index (CFI) and the Tucker Lewis Index (TLI) assess to what extent a model reproduces the observed covariance matrix better than a baseline model that is assuming that all observed variables are uncorrelated (estimates of $> .90$ point to an acceptable fit and $> .95$ to a good fit).

Whether M_JOY and M_SELF were the same in Norway one week before, on the examination day, and 3–5 weeks after the examination, as well as the same between Norway and Germany, was tested via examination of measurement invariance. A bottom-up approach was applied by examining configural invariance first, which means that the same items had to be associated with M_JOY and M_SELF at each time point or in each country. Second, metric invariance was tested for, which means that the factor loadings of the items used to assess the two constructs were restricted to be the same across time points and countries. The third step was a test for scalar invariance, which means that the item intercepts were restricted to be the same (Rutkowski & Svetina, 2014). The fit of the different models was then compared.

The development of preschool teachers' attitudes in Norway from before to after their examinations was estimated by applying linear latent growth modeling. The measures of M_JOY and M_SELF at the three time points served as indicators of the respective growth factors. Random effects were specified to capture individual differences in development.

Finally, we applied latent class growth analysis (LCGA) to find out whether groups of prospective preschool teachers with different latent trajectories existed in Norway and whether the last mathematics grade was able to predict class membership (McLachlan & Peel, 2000). To make the models not overly complicated, no variation within these groups was allowed (Kreuter & Muthén, 2008). Part of the LCGA was a decision about the number of classes necessary to describe the development of prospective preschool teachers across time points. This was done by comparing models with one, two, three, or four classes, respectively. We evaluated their classification quality based on an aggregated uncertainty measure called “entropy” (Ramaswamy, DeSarbo, Reibstein, & Robinson, 1993),

whereby estimates close to one indicate well-separated classes and estimates close to zero indicate ill-fitting models. In addition, the estimate of the mean probability for prospective preschool teachers' most likely latent class membership was taken into account. Finally, we evaluated relative fit criteria (Nylund, Asparouhov, & Muthén, 2007), whereby lower absolute values indicate a better-fitting model.

All analyses were carried out using Mplus 7.3 (Muthén & Muthén, 2014). A robust maximum-likelihood estimator (Sass, Schmitt, & Marsh, 2014) and a sandwich-type covariance matrix were applied to achieve precise estimations of standard errors (Satorra & Bentler, 2001) and chi-square statistics robust to non-normality of the data (Yuan & Bentler, 2000). Full Information Maximum Likelihood (FIML) estimation, integrating missing data analyses and parameter estimation under the missing at random assumption, was used to handle partially missing data (Little & Rubin, 2014).

Results

Stability of M_JOY and M_SELF across Time Points (H1)

The test on longitudinal measurement invariance revealed in the case of M_JOY that the metric model did not fit significantly less well to the Norwegian data than the configural model (see Table 3). As hypothesized, this indicated stability in the sense that the different items were associated with M_JOY in a similar way no matter whether prospective preschool teachers were working on the survey one week before their examination, on the examination day itself, or had received their grades 3–5 weeks later. In contrast, the fit decreased significantly when the intercepts were fixed to be the same across the three time points. This result indicated variability of the initial level of JOY. An item-by-item inspection revealed that the difference in intercepts was largest in the case of the items “Mathematics leads to enjoyable experiences” and “Mathematics is enjoyable” at the third time point, where the intercepts were significantly higher than at the other two time points.

Variability was larger with respect to M_SELF because there was already a significant difference between the configural and metric models, indicating that the items were associated with M_SELF in different ways at the different time points. The variability in factor loadings was related to the item “I know different ways to solve difficult mathematical problems.” The loading was significantly higher at the third time point.

Stability of M_JOY and M_SELF across Countries (H2)

The test on measurement invariance across countries revealed stability as hypothesized, in the sense that configural and metric but not scalar invariance existed in the case of M_JOY (see Table 4). All intercepts were significantly higher in Norway than in Germany. In the case of M_SELF, metric invariance was not supported by the data. It was the item “Although it is exhausting, I am able to push myself and to accomplish high achievements in mathematics” that revealed the biggest difference. It was much stronger related to M_SELF in Norway than in Germany.

Table 3. Results of the tests on longitudinal measurement invariance.

	Mathematics-related enjoyment		Mathematics-related self-efficacy	
	Chi-square	df	Chi-square	df
Configural model	221.8	83	134.3	51
Metric model	224.0	91	146.8	57
Scalar model	242.4	99	152.7	63
Difference _{Con_Met}	−2.2	8	−12.5*	6
Difference _{MET_SCA}	−18.4*	8	−5.9	6

Note: df = degrees of freedom.

Table 4. Results of the tests on cross-country measurement invariance.

	Mathematics-related enjoyment		Mathematics-related self-efficacy	
	Chi-square	df	Chi-square	df
Configural model	41.6	10	3.4	4
Metric model	48.2	14	14.5	7
Scalar model	141.2	18	33.7	10
Difference _{Con_Met}	-6.6	4	-11.1*	3
Difference _{MET_SCA}	-93.0*	4	-19.2*	3

Note: df = degrees of freedom.

Prior Mathematical Knowledge as a Predictor of Attitudes in Norway and Germany (H3)

M_JOY was relatively high among prospective preschool teachers at the initial time point (intercept [I] = 4.27; standard error [SE] = .07 on a scale from 1–6). It remained stable over the course of the following five weeks (slope [S] = .01; SE = .04). M_SELF started relatively low (I = 1.82; SE = .04 on a scale from 1–4) but grew significantly over time (S = .09*; SE = .04). An inspection of the piecewise developments between time points 1 and 2 as well as between 2 and 3 revealed no significant differences between these periods in the case of M_JOY. In the case of M_SELF, growth between the second and third time points was higher than between the first and second time points ($S_1 = .07^\dagger$, SE = .04; $S_2 = .24^*$, SE = .13).

Prior knowledge in mathematics was as hypothesized, not only a significant predictor of the initial level of prospective preschool teachers' M_JOY and M_SELF in Norway but also of their developmental trajectories. The better was the mathematics school grade, the more Norwegian student teachers enjoyed mathematics at the beginning of the measurements (I_{M_Joy} on MCK = .33*, SE = .10) and the higher was their M_SELF (I_{M_Self} on MCK = .20*, SE = .03). Furthermore, the better was the school grade in mathematics, the more positively student teachers' M_JOY tended to develop over time (S_{M_Joy} on MCK = .07[†], SE = .04) and the more positively M_SELF developed (S_{M_Self} on MCK = .11*, SE = .03).

In Germany, the relationship of initial M_JOY and M_SELF to the grade in school mathematics was even stronger: $b = -.49$ (.05)* in the case of M_JOY and $b = -.40$ (.03)* in the case of M_SELF (note that the grade scale in Germany is in reverse to that applied in Norway).

Differential Trajectories of Preschool Teachers in Norway?

Based on the information criterion from the LCGA (see Table 5), a distinction between two trajectories should be made because the fit improved in this case whereas improvement in fit was marginal in cases with three or four classes. This pattern applied both to M_JOY and M_SELF. Entropy parameters and classification probabilities were sufficient in two- and three-class models. Substantive considerations favored the three-class over the two-class solution, both with respect to M_JOY and M_SELF, because three clearly distinct trajectories could be identified then. The

Table 5. Results of the latent class growth analyses (LCGA) by construct.

# cl.	M_JOY				M_SELF			
	Count estimates	Entropy	Classification probability	BIC _{adj}	Count estimates	Entropy	Classification probability	BIC _{adj}
1	221	1.00	1.00	888.6	221	1.00	1.00	569.1
2	203/18	.81	.96/.82	884.7	175/46	.75	.95/.85	549.4
3	185/18/17	.80	.93/.83/.74	886.6	166/45/9	.83	.95/.85/.83	534.1
4	164/33/18/7	.66	.87/.62/.70/.79	886.3	125/42/40/14	.70	.89/.63/.79/.91	539.9

Note: BIC_{adj} = adjusted Bayesian Information Criterion.

four-class solution did not add new information in terms of differential trajectories but distinguished only between two groups with slightly higher and lower estimates. We decided therefore to apply the three-class solution.

In the case of M_JOY, by far the largest of the three groups identified included 185 prospective preschool teachers who slightly disfavored mathematics, and this reluctant attitude remained stable over time. It was not affected by the examination day or by receiving their grades 3–5 weeks later ($I = 3.07$ [.30]*, $S = .12$ [.19]). In contrast, one small group of 18 preschool teachers existed who had a very low M_JOY level and even this decreased significantly over time ($I = 1.94$ [.38]*, $S = -1.00$ [.40]*). A third small group of 17 members had a significantly higher M_JOY level, above the neutral midpoint of the scale, in the beginning and this level remained stable over the course of the five weeks with the examination in between ($I = 4.19$ [.30]*, $S = -0.17$ [.16]).

In the case of M_SELF, the largest of the three groups identified included 166 prospective preschool teachers who had the lowest self-efficacy at the beginning and this did not change throughout the three time points ($I = 1.08$ [.13]*, $S = -0.04$ [.11]). In addition, a relatively large group of 45 preschool teachers existed who reported significantly higher M_SELF at the beginning. However, its level was still below the neutral midpoint of the scale, and it decreased significantly over the course of three time points with the examination in between ($I = 2.04$ [.21]*, $S = -0.39$ [.18]*). Only the M_SELF level of a remaining small group with nine members was relatively high, namely, at around the neutral midpoint of the scale, and it remained stable between the three time points ($I = 2.74$ [.21]*, $S = 0.03$ [.20]).

A combined LCGA that included both constructs at the same time revealed that their initial levels were correlated ($I_{M_Joy} * I_{M_Self} = .33^*$, $SE = .16$; standardized estimates). Good entropy (.71), classification probabilities (.86/.77/.84/.91) and substantive considerations led in this case to a solution with two classes for each construct (see Table 6), although models with more classes had sufficient entropy and classification probability, too. However, the larger number of class patterns was almost impossible to interpret and tended to split the sample up into very small groups with only two to five members.

We find two groups each with initial M_JOY slightly below (classes 1 and 3) or slightly above the neutral midpoint of the scale (classes 2 and 4). Each of the two groups is then split into one group with low (classes 1 and 2) and one group with slightly higher (classes 3 and 4) M_SELF. The two groups with higher M_JOY have in each case higher M_SELF estimates than the two groups with lower M_JOY, and M_SELF in each case remains stable. M_JOY either remains stable (class 2; marked green in Figure 1) or decreases over time in the case of the highest starting level (class 4; marked blue). These two groups, together, comprise 30% of the sample.

In contrast, M_SELF decreases significantly over time in both classes with lower M_JOY, although M_SELF is already at a low level from the beginning. In addition, M_JOY tends to decrease even further in case of the class with the lowest M_JOY level at the beginning (class 1; marked red) whereas it remains stable in the case of the class with a low but slightly higher level of M_JOY (class 3; marked orange). Class 1 includes half of the sample, and class 2 includes about 20%, so that, overall, they comprise 70% of the Norwegian prospective preschool teachers.

Table 6. Results of the combined latent class growth analysis (LCGA).

Class	I_{M_Joy}	S_{M_Joy}	I_{M_Self}	S_{M_Self}	Proportion (%)
1	2.70 (.35)*	-.24 (.13) [†]	0.77 (.11)*	-.32 (.11)*	49
2	3.81 (.28)*	-.01 (.17)	1.14 (.13)*	.07 (.14)	25
3	3.28 (.44)*	-.20 (.18)	1.75 (.15)*	-.61 (.15)*	21
4	3.98 (.38)*	-.68 (.19)*	2.53 (.15)*	.13 (.17)	6

Note: I = intercept, S = slope.



Figure 1. Development of mathematics-related enjoyment and self-efficacy over time.

Discussion

From a longitudinal perspective across very different situations (with and without an examination), it was possible to confirm stability in terms of configural and metric invariance with respect to M_JOY over three time points but not with respect to M_SELF. This means that, in the latter case, variability existed and the measurement requirement for comparing relations across time points was not met. Results should therefore be used with caution.

The same pattern applied to the cross-country comparisons. It was possible to confirm configural and metric invariance with respect to M_JOY but not with respect to M_SELF. These results may reflect the unclear state of research with respect to the stability of attitudes, which were sometimes hypothesized to be fully determined by the situation (Wilson & Hodges, 1992) and at other times to be rather stable. Both extremes may not be appropriate, however, because our results suggest that it may not be possible to generalize research results across psychological constructs. A general state of research “on attitudes” would not exist then because results would depend on the specific construct examined. In our case, the study points to the larger stability of M_JOY across different contexts than M_SELF.

With respect to the sample of Norwegian prospective preschool teachers, M_SELF increased significantly before and in particular after the examination, which may be interpreted as a result of increased learning time invested in mathematics due to the examination and the actual experience of being able to solve mathematical problems. The literature suggests that learning time is generally positively related to achievement (e.g., Keeves, 1992), which in turn is positively related to self-efficacy (Pajares, 1997). This may also apply in this case.

The interesting phenomenon in this context is that the increase occurred without a decrease in M_JOY. One might have expected that practicing for the examination was a burden for the student teachers, which resulted in a loss of motivation. This is a popular and recurring theme in the media, but we did not find evidence for this—at least not as long as we examined the full sample of prospective preschool teachers.

Exploratory analyses revealed that different subgroups existed in the Norwegian sample of prospective preschool teachers, with a worryingly large proportion (about half of the sample) reporting low enjoyment and low self-efficacy, including the negative developments of such. An important result of our study may therefore be that initial M_JOY and M_SELF levels, as well as the differential trajectories, could be predicted by students’ final school grades in mathematics. Assuming that these reflect prior knowledge, it may be concluded that these grades are able to protect against unfavorable attitudes towards mathematics. This conclusion is supported by data from another country,

Germany, in terms of the strong relationship between school grades in mathematics to M_JOY and M_SELF.

Limitations

Before we turn to conclusions, several limitations of our study have to be pointed out that may have affected the results. First, the drop-out rate after the examination was high. Although bias with respect to prior knowledge or gender did not exist, only a small, selected group of preschool teachers took part in the third measurement. An item-by-item inspection of the M_JOY scale revealed that the intercepts of two items were significantly higher at the third time point than at the other two time points. Growth in M_SELF was also higher between the second and the third time points compared to the period between the first two time points. These differences may indicate that the smaller sample was nevertheless positively biased and comprised particularly motivated students. Including the third time point can thus serve heuristic purposes only.

Second, we examined a selected set of prospective preschool teachers' characteristics only. Other characteristics, for example their enjoyment of and self-efficacy in respect to actually *implementing* mathematical activities—not mathematics as merely a subject—may affect teachers' behavior in preschool, too. Such reflections are important because Gregoire (2003) showed, in her systematic review of research on conceptual change, that different types of beliefs or attitudes may conflict with each other and, thus, block the implementation of related behavior. A typical case in our context could be that preschool teachers value certain types of implementation strategies and activities with children but hold negative attitudes toward mathematics. How teachers act in this case remains an open question. It is therefore important to identify such conflicts in further studies so that it may be possible to implement interventions to address them.

Conclusions

With respect to the examination, it transpired that, on average, negative effects were limited. However, some subgroups experienced negative developments before and after the examination. Given that we probably used a positively biased sample, the problem may be larger in reality than documented; it is thus worthwhile to examine such effects in more detail. One research question that could be posed, for example, is whether different placements or types of examination (earlier, later; several smaller examinations rather than a single large one) within the teacher education program matter.

With the sensitive population of preschool teachers in mind who generally seem to dislike mathematics, it may also be worthwhile to think carefully through the sequence of studies during teacher education, in particular where to place mathematical experiences. Fostering early mathematical literacy is nowadays an inherent part of a preschool teachers' work. However, as a result of the late evolvement of corresponding awareness, many open questions exist that require more research attention in the near future.

Our research intended to create a starting point for this. M_JOY and M_SELF are important antecedents for implementing mathematics-related activities in preschool and, thus, for fostering the development of children's early mathematical literacy. If this is what preschool teacher education and policy makers are interested in, and Norway is a country where this has recently come to attention (Mosvold & Sæbbe, 2015), ways to improve M_JOY and M_SELF should be sought. Our results point to two potential conclusions. Since prior school grades predict higher levels of these constructs and more positive developments, selection of candidates for preschool teacher education according to this criterion is a possibility. However, given the low average grade in mathematics of current applicants, this could result in a serious lack of students.

Another means of improving the situation may therefore be more promising in the long run: investing time during preschool teacher education in providing prospective teachers themselves

with enjoyable and successful experiences of mathematical activities. Examples include creating a mathematics laboratory or holding mathematics days during which both young children and prospective preschool teachers experience mathematics in a varied, creative, and joyful way. Student teachers could then reflect on these experiences during their classes in mathematics education (Thiel, 2015).

Such an approach allows for increasing knowledge while engaging in discussions and reflections (Gregoire, 2003) and may have the potential to create positive attitudes or challenge existing negative attitudes. To our knowledge, only one teacher education program in Norway has systematically implemented such a strategy (<http://dmmh.no/for-barnehagene/matematikkrommet>), which is intended to facilitate the interaction of children and prospective preschool teachers through fostering children's early mathematical literacy (Nakken & Thiel, 2014). One can reasonably assume that M_JOY and M_SELF can be strengthened in this way.

Disclosure Statement

No potential conflict of interest was reported by the authors.

References

- Alvestad, M., Johansson, J.-E., Moser, T., & Søbstad, F. (2009). Status og utfordringer i norsk barnehageforskning. *Nordic Early Childhood Education Research*, 2, 39–55.
- Alvestad, M., & Samuelsson, I. P. (1999). A comparison of the national preschool curricula in Norway and Sweden. *Early Childhood Research and Practice*, 1, Retrieved from <http://ecrp.uiuc.edu/v1n2/alvestad.html>
- Anastasi, A. (1983). Traits, states, and situations: A comprehensive view. In H. Wainer, & S. Messick (Eds.), *Principles of modern psychological measurement* (pp. 345–356). Hillsdale, NJ: Erlbaum.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191–215.
- Bates, A. B., Latham, N., & Kim, J. (2011). Linking preservice teachers' mathematics self-efficacy and mathematics teaching efficacy to their mathematical performance. *School Science and Mathematics*, 111(7), 325–333.
- Blömeke, S., Dunekacke, S., & Jenßen, L. (2017). Cognitive, educational and psychological determinants of prospective preschool teachers' beliefs. *European Early Childhood Education Research Journal*, 25(6), 885–903.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, 43, 1428–1446.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An Introduction to theory and research*. Reading, MA: Addison-Wesley.
- Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. *Social Policy Report: Giving Child and Youth Development Knowledge Away*, 22(1), 1–23.
- Gregoire, M. (2003). Is it a challenge or a threat? A dual-process model of teachers' cognition and appraisal processes during conceptual change. *Educational Psychology Review*, 15, 147–179.
- Guo, Y., Piasta, S. B., Justice, L. M., & Kaderavek, J. N. (2010). Relations among preschool teachers' self-efficacy, classroom quality, and children's language and literacy gains. *Teaching and Teacher Education*, 26, 1094–1103.
- Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy/mathematics performance correspondence. *Journal for Research in Mathematics Education*, 20, 261–273.
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18, 117–144.
- Hu, L., & 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.
- Jenßen, L., Dunekacke, S., Eid, M., & Blömeke, S. (2015). The relationship of mathematical competence and mathematics anxiety: An application of latent state-trait theory. *Zeitschrift für Psychologie*, 223, 31–38.
- Keeves, J. P. (1992). *The IEA study of science: Changes in science education and achievement: 1970 to 1984*. Oxford: Pergamon.
- Kowalski, K., Pretti-Frontczak, K., & Johnson, L. (2001). Preschool teachers' beliefs concerning the importance of various developmental skills and abilities. *Journal of Research in Childhood Education*, 16, 5–14.
- Krajewski, K., & Schneider, W. (2009). Exploring the impact of phonological awareness, visual-spatial working memory, and preschool quantity-number competencies on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of Experimental Child Psychology*, 103, 516–531.

- Kreuter, F., & Muthén, B. (2008). Analyzing criminal trajectory profiles: Bridging multilevel and group-based approaches using growth mixture modeling. *Journal of Quantitative Criminology*, 24(1), 1–31.
- Laschke, Ch., & Blömeke, S. (2016). Measurement of job motivation in TEDS-M: Testing for invariance across countries and cultures. *Large-Scale Assessments in Education*, 4, 16. doi:10.1186/s40536-016-0031-5
- Lee, J. S., & Ginsburg, H. P. (2007). Preschool teachers' beliefs about appropriate early literacy and mathematics education for low- and middle-socioeconomic status children. *Early Education and Development*, 18(1), 111–143.
- Liebert, R. M., & Liebert, L. L. (1998). *Liebert & Spiegler's personality: Strategies and issues* (8th ed.). Pacific Grove, CA: Brooks/Cole.
- Little, R. J., & Rubin, D. B. (2014). *Statistical analysis with missing data*. Hoboken, NJ: John Wiley & Sons.
- McCrae, R. R., & Costa, P. T., Jr. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 51–87). New York, NY: Guilford Press.
- McLachlan, G. J., & Peel, D. (2000). *Finite mixture models*. New York, NY: Wiley.
- McMullan, M., Jones, R., & Lea, S. (2012). Math anxiety, self-efficacy, and ability in British undergraduate nursing students. *Research in Nursing & Health*, 35, 178–186.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525–543.
- Mosvold, R. (2012). Førskolelærerens utfordringer knyttet til arbeidet med antall, rom og form i barnehagen. In M. Alvestad, & T. Vist (Eds.), *Læringskulturer i barnehagen: Flerfaglige forskningsperspektiver* (pp. 92–114). Oslo: Cappelen Damm Akademisk.
- Mosvold, R., Bjuland, R., Fauskanger, J., & Jakobsen, A. (2011). Similar but different – Investigating the use of MKT in a Norwegian kindergarten setting. In M. Pytlak, T. Rowland, & E. Swoboda (Eds.), *Proceedings of the seventh congress of the european society for research in mathematics education* (pp. 1802–1811). Rzeszów: University of Rzeszów.
- Mosvold, R., & Sæbbe, P.-E. (2015). Tasks of teaching mathematics in kindergarten. In H. Silfverberg, T. Kärki, & M. S. Hannula (Eds.), *Nordic research in mathematics education – Proceedings of NORMA14, Turku, June 3-6, 2014* (pp. 371–380). Studies in Subject Didactics 10. Turku: The Finnish Research Association for Subject Didactics.
- Muthén, L. K. & Muthén, B. O. (2014). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Nakken, A., & Thiel, O. (2014). *Matematikkens kjerne*. Bergen: Fagbokforlaget.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(4), 535–569.
- OECD. (2006). *Starting strong II: Early childhood education and care*. Paris: OECD publishing.
- Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10; pp. 1–49). Greenwich, CT: JAI Press.
- Ramaswamy, V., DeSarbo, W., Reibstein, D., & Robinson, W. (1993). An empirical pooling approach for estimating marketing mix Elasticities with PIMS data. *Marketing Science*, 12, 103–124.
- Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula (Ed.), *Handbook of research on teacher education* (2nd ed., pp. 102–119). New York: Macmillan.
- Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, 74(1), 31–57.
- Sass, D., Schmitt, T., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling: A Multidisciplinary Journal*, 21, 167–180.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507–514.
- Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35–37). Windsor: NFER-NELSON.
- Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest and academic engagement. *The Journal of Educational Research*, 95, 323–332.
- Sommersel, H. B., Vestergaard, S., & Larsen, M. S. (2013). *Kvalitet i barnehager i skandinavisk forskning 2006–2011*. København: Dansk Clearinghouse for Uddannelsesforskning.
- Spielberger, C. D., & Reheiser, E. C. (2009). Assessment of emotions: Anxiety, anger, depression, and curiosity. *Applied Psychology: Health and Well-Being*, 1, 271–302.
- Statistisches Bundesamt. (2014). *Statistiken der Kinder- und Jugendhilfe: Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2014*. Wiesbaden: Statistisches Bundesamt.
- Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., ... Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Findings from the IEA teacher education and development study in mathematics (TEDS-M)*. Amsterdam: IEA.
- Thiel, O. (2015). *A room for mathematics in early childhood*. Paper presented on September 9th, 2015 at 25th EECERA Annual Conference, September 7th till 10th, 2015 in Barcelona, Catalunya, Spain.

- Tietze, W., Roßbach, H.-G., & Grenner, K. (2005). *Kinder von 4 bis 8 Jahren: Zur Qualität der Erziehung und Bildung in Kindergarten, Grundschule und Familie*. Weinheim: Beltz.
- Tsamir, P., Tirosh, D., Levenson, E., Tabach, M., & Barkai, R. (2015). Preschool teachers' knowledge and self-efficacy needed for teaching geometry: Are they related? In B. Pepin & B. Roesken-Winter (Eds.), *From beliefs to dynamic affect systems in mathematics education* (pp. 319–337). Heidelberg: Springer International Publishing.
- Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Newbury Park, CA: Sage.
- Van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, 54, 119–135.
- Varol, F. (2013). What they believe and what they do. *European Early Childhood Education Research Journal*, 21(4), 541–552.
- Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 37–65). Hillsdale, NJ: Erlbaum.
- Yuan, K., & Bentler, P. M. (2000). 5. Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, 30(1), 165–200.