# Assessing the psychometric properties of the Behavior and Instructional Management Scale: a study on a sample of Serbian teachers

Stefan Ninković & Olivera Knežević Florić

*Department of Pedagogy, Faculty of Philosophy, University of Novi Sad, Serbia*

The Behavior and Instructional Management Scale (BIMS) was developed to assess the bidimensional construct of classroom management. The purpose of this study was to evaluate the factor structure and psychometric properties of the BIMS using Serbian teachers. Confirmatory factor analysis was conducted on the data collected from a sample of 660 teachers, with results supporting a two-factor model of the BIMS. Both subscales of the BIMS demonstrated adequate internal consistency. Furthermore, results indicated that the two-factor model of the scale has good convergent validity. In conclusion, the BIMS can be recommended to researchers interested in measuring teacher classroom management in Serbia.

**Key words:** behavior management, instructional management, teachers, classroom management, confirmatory factor analysis, Serbia

Highlights:

- A two-factor 12-item model of the BIMS shows good fit indices.
- Findings provide support for the reliability and convergent validity of the BIMS.
- Results indicate significant differences with teacher gender and grade level.

Numerous studies have converged on the conclusion that teacher quality is an important determinant of students' achievement (Darling-Hammond & Youngs, 2007; Hattie, 2009). High-quality teaching is hard to achieve without effective classroom management, which is widely conceptualized as "the actions teachers take to create an environment that supports and facilitates both academic and social-emotional learning" (Evertson & Weinstein, 2006, p. 4). The two basic purposes of classroom management are to create opportunities for student learning and to prevent disruptive behavior (Aloe, Amo, & Shanahan, 2014; Korpershoek, Harms, de Boer, van Kuijk, & Doolaard, 2016).

Research evidence indicates that components of effective classroom management, such as clear rules of behavior, consistent routines and efficient use of time, significantly contribute to the improvement of student achievement (Stronge, Ward, & Grant, 2011). In a recent meta-analysis, Korpershoek et al. (2016) have found that successful classroom management reduces behavioral problems and improves students' academic achievements. In contrast, by using ineffective classroom management strategies, teachers aggravate the classroom climate, which can result in more frequent student misbehaviors and increased stress for both students and teachers (Oberle & Schonert-Reichl, 2016).

## The Behavior and Instructional Management Scale (BIMS)

Since scholars have conceptualized classroom management in different ways, it is not surprising that there are multiple measures of this construct. However, some instruments that have been used in previous research have various conceptual and measurement issues. For example, the Attitudes and Beliefs on Classroom Control Inventory (ABCC) (Martin, Yin, & Mayall, 2008) is characterized by high inter-correlation of latent factors, thereby compromising the discriminant validity of the instrument. Other research instruments that have been utilized in some studies (Nie & Lau, 2009) include measures such as behavioral control and care, but they lack an important dimension of instructional management. Finally, it should be noted that scientists have devoted attention to the investigation of classroom management self-efficacy. However, this construct refers to teacher's perceived ability to maintain classroom order and control, but does not include actual teaching behaviors (O'Neill & Stephenson, 2011).

Considering key conceptual and measurement issues surrounding research on classroom management, Martin & Sass (2010) developed the Behavior and Instructional Management Scale (BIMS). This measure is based on conceiving classroom management as a bifaceted construct that includes two dimensions: behavior management and instructional management. Behavior management refers not only to pre-planned activities of teachers that aim to prevent disciplinary problems, but also to respond to them. Instructional management comprises of classroom practices that involve the establishment of everyday work routine, fostering interaction between students, and enabling students to take an active role in learning activities (Martin & Sass, 2010). The authors of this instrument assume that teachers can have incongruent classroom management beliefs. This

means that teachers who have constructivist beliefs about the nature of student learning can at the same time have more rigid views about classroom behavior rules.

The initial validation of the BIMS was carried out on a sample of 550 teachers from the United States (Martin & Sass, 2010). Using the exploratory factor analysis, the initial 24-item version of the scale was reduced to a version with 12 items that best measure the hypothesized latent construct. Two extracted latent factors were interpreted as behavior management (BM) and instructional management (IM). Confirmatory factor analysis also supported this two-factor solution. Later research (Sass, 2011) demonstrated that the IM subscale is invariant across the grade level, while the BM subscale was only partially invariant.

Psychometric characteristics of the BIMS were also evaluated on a sample of Portuguese teachers (Sass, Lopes, Oliveira, & Martin, 2016). Although the two-factor structure of the instrument was confirmed, it was found that two items from the original 12-item version of the scale have high cross loadings. Therefore, these two items, based on the results of the exploratory factor analysis of the 24-item version, were replaced with another two items. Scores on two subscales were significant predictors of teacher's self-efficacy in the area of student engagement, instructional strategies and classroom management.

There are sources suggesting that the BIMS was translated into Turkish (Ünal & Ünal, 2012) and Iranian language (Jalali, Panahzade, & Firouzmand, 2014). However, information regarding the structural validity of these translations is not provided in the literature accessible to authors of this paper. Bearing in mind the previous studies, it is evident that the BIMS is a relatively new instrument which requires further research into its psychometric characteristics in other cultural and educational contexts.

Convergent validity of the BIMS has been examined by exploring its associations with other relevant constructs. For example, in a study conducted on a sample of teachers from the United States (Martin & Sass, 2010), the relations between scores on the BM and IM subscales and teacher self-efficacy were examined. The concept of teacher self-efficacy is defined as teachers' beliefs about their abilities to influence students' learning and involves three components: efficacy for student engagement, efficacy for classroom management, and efficacy for instructional strategies (Tschannen-Moran & Woolfolk Hoy, 2001). Martin and Sass (2010) reported a low negative correlation between BM and efficacy for classroom management. The IM has been found to have strong negative correlations with all three dimensions of teacher self-efficacy. These findings suggest that high levels of teacher self-efficacy are associated with flexible, constructivist approaches to instruction.

Results from empirical studies have shown that teacher's responses to measures of classroom management can vary due to demographic and job-related characteristics. Some studies have demonstrated that male teachers are more in control of instruction than female teachers (Lam, Tse, Lam, & Loh,

2010; Martin, Sass, & Schmitt, 2012). The teaching level has also been linked to classroom management strategies. In particular, empirical evidence suggests that teachers working in elementary schools show significantly less control over instructional activities than middle and high school teachers (Martin et al., 2012). With regards to teacher experience, previous studies have shown that more experienced teachers score higher on the instructional management subscale (Martin, Yin, & Mayall, 2006).

## Purpose of the Study

In order to ensure the validity of the research of the classroom management in Serbia, researchers need psychometrically sound measures of this construct. The BIMS is a promising measure of behavior and instructional management, two core dimensions of classroom management. Therefore, the first aim of this study was to evaluate the factor structure of the BIMS in a sample of the Serbian teachers. We expected that a latent factor model of the scale with two dimensions would best fit the data obtained from the sample of Serbian teachers. Our second aim was to examine the reliability and convergent validity of the Serbian version of the BIMS. We hypothesized that the Serbian version of the BIMS would have good internal consistency reliability. In line with what has been found with teachers in the United States (Martin & Sass, 2010), we expected that scores on the IM subscale would be negatively correlated with the dimensions of teacher self-efficacy, and that the BM would correlate negatively with teacher efficacy for classroom management.

## Method

### Sample and Procedure

The final sample consisted of 660 teachers (78.1% females) who work in elementary schools (57.6%) and high schools (42.4%) in Serbia. In Serbia, teachers work on three levels of the educational system: elementary school classroom teaching (grades 1–4), secondary school subject teaching (grades 5–8), and high school teaching (grades 9–12). The average teacher's age was 44.13 ($SD = 9.14$) years. Table 1 shows the participants' descriptive characteristics that were examined in line with Sass et al. (2016).

School principals' approvals for the research were obtained prior to teachers' participation in the study. After obtaining the permission, one of the authors of this article visited the schools and administered the survey at suitable times. In twelve schools that were visited by one of the authors of this paper, the data was collected using a standard paper-and-pencil procedure. A subsample of teachers ($N = 283$) filled the instruments using an online survey. This allowed the teachers from schools across the Serbia to participate in the study. All teachers had been informed that participation in the study was anonymous and voluntary, and that they could withdraw from the study at any moment.

Table 1
*Participants' socio-demographic characteristics*

|  | Frequency | Percentage (%) |
|---|---|---|
| Gender | | |
| Male | 144 | 21.88 |
| Female | 514 | 78.12 |
| Teaching experience | | |
| 1 year | 26 | 3.94 |
| 2–4 years | 47 | 7.12 |
| 5–9 years | 106 | 16.06 |
| 10–15 years | 125 | 18.94 |
| More than 15 years | 356 | 53.94 |
| School level | | |
| Primary elementary education | 173 | 26.41 |
| Secondary elementary education | 204 | 31.15 |
| Grammar and vocational school | 278 | 42.44 |
| How much time during class do you spend in managing student behavior? | | |
| Less than 10% | 132 | 20.18 |
| 10%–20% | 179 | 27.37 |
| 20%–30% | 105 | 16.05 |
| 30%–40% | 71 | 10.86 |
| 40%–50% | 46 | 7.03 |
| 50%–60% | 42 | 6.42 |
| 60%–70% | 21 | 3.21 |
| 70%–80% | 32 | 4.89 |
| 80%–90% | 26 | 3.98 |
| On average, how often do you report students' misbehavior to others at school? | | |
| Rarely | 291 | 44.16 |
| A few times a year | 67 | 10.17 |
| Sometimes | 253 | 38.39 |
| Often | 48 | 7.28 |
| Have you participated in professional training programs in the field of classroom management? | | |
| No | 236 | 35.81 |
| 1–8 hours | 197 | 29.89 |
| 9–16 hours | 107 | 16.24 |
| 17–24 hours | 52 | 7.89 |
| Over 24 hours | 67 | 10.17 |
| Based on your experience, in the last five years student misbehavior has been... | | |
| Significantly reduced | 16 | 2.44 |
| Reduced | 33 | 5.04 |
| Has not changed | 174 | 26.56 |
| More frequent | 288 | 43.97 |
| Significantly more frequent | 144 | 21.98 |

*Note.* As a result of the missing data, the sum of answers on certain questions is not 660.

## Instruments

**The Behavior and Instructional Management Scale.** The Behavior and Instructional Management Scale (Martin & Sass, 2010) contains two subscales: behavior management (BM) and instructional management (IM). The original version of the scale contains 12 items, which are equally distributed in the two subscales. However, keeping in mind the results of previous research (Sass et al., 2016) and the recommendations of the authors of the scale, we decided to use the 14-item version of the instrument in order to determine which 12-item version of the BIMS has the best psychometric properties in the Serbian sample. All items were scored on a 6-point Likert scale (1 = *strongly disagree*, 6 = *strongly agree*). High scores on subscales indicated the teacher's more controlling approach to student behavior management and instruction.

We used a back-translation procedure to translate the items of the English version of the BIMS. First, one author of the present study translated the BIMS into the target language, Serbian. Then two experienced teachers, independent of one another, read all items and evaluated that they are relevant to the educational context in Serbia. The instrument was then translated back to English by a professional translator. The translator had no access to the original English version of the instrument. Finally, the researchers compared the back-translated Serbian version of the BIMS and the original version of the scale. Some minor discrepancies that were found between the two versions were resolved. The Serbian translation of the BIMS is available from the corresponding author.

**The Teachers' Sense of Efficacy Scale.** In order to examine the convergent validity of the BIMS, the Serbian version of Teachers' Sense of Efficacy Scale – TSES was applied. Specifically, the short form of the TSES (Tschannen-Moran & Woolfolk Hoy, 2001) which was validated in Serbia by Ninković and Knežević Florić (2018) was used. The TSES-short version contains 12 items distributed in 3 subscales: efficacy for student engagement (e.g., "How much can you do to help your students value learning?"), efficacy for classroom management (e.g., "How much can you do to control disruptive behavior in the classroom?"), and efficacy for instructional strategies (e.g., "To what extent can you provide an alternative explanation or example when students are confused?"). A 9-point Likert scale was used for all items (1 = *none at all*, 9 = *a great deal*). In the present study, reliability coefficients for the Serbian version of the TSES were high: efficacy for student engagement ($\alpha$ = .76), efficacy for classroom management ($\alpha$ = 0.88), efficacy for instructional strategies ($\alpha$ = .77).

## Statistical Analysis

The BIMS structure was examined using the confirmatory factor analysis (CFA). Since the chi-square test of exact fit is usually statistically significant in studies with large samples, we relied on the following goodness-of-fit indices: the comparative fit index (CFI), the Tucker-Lewis Index (TLI), and root mean square error of approximation (RMSEA). For the CFI and TLI values larger than .95 are considered to represent a good model fit, and values between .90 and .95 are interpreted as indicators of an acceptable fit; for the RMSEA values smaller than .06 point to a good model fit, and values between .06 and .08 indicate an acceptable fit to the data (Hu & Bentler, 1999). We also reported the weighted root mean square residual (WRMR) fit statistic. Although the WRMR values of about 1.00 are considered acceptable, it should be noted that simulation studies have shown that the WRMR might provide misleading results if large samples are used (DiStefano, Liu, Jiang, & Shi, 2018). The parameter estimates were obtained using WLSMV estimator, an estimation method which was used in previous validation studies (Martin & Sass, 2010; Sass et al., 2016).

After the factor structure of the BIMS had been examined, internal consistency coefficients for scores on the BIMS were calculated. Pearson's product-moment correlation coefficient was used to evaluate the convergent validity of the scale. Finally, differences in the scores on the subscales of the BIMS were analysed using a multivariate analysis of variance (MANOVA).

The missing data was handled with package mice in R software (van Buuren, 2018). The missing data was imputed in two steps. First, teachers who did not respond to more than 5% of the questions from the given survey were excluded from further analysis, which resulted in a sample of 660 teachers. After that, missing data was imputed using the method of multiple imputations (MI) by chained equations. We generated five completed datasets and subsequently did all analysis using the first imputed dataset. SPSS version 23 was used for conducting descriptive statistics. Both Cronbach's alpha and omega coefficients of internal consistency were estimated using a polychoric correlation matrix in the userfriendlyscience R package (Peters, 2014). The confirmatory factor analysis was performed in Mplus 7.31 (Muthén & Muthén, 2015).

## Results

Table 1 shows the descriptive statistics of the fourteen BIMS items, as well as the average subscale scores for the BM and IM. As it can be seen, the skewness values for all items were between −1.29 and 1.21, while the kurtosis values were in the range from −0.22 to 2.20. Mean value on the subscale BM was significantly higher than on the subscale IM, $t(659) = 56.50$, $p < .001$, $d = 2.20$. Based on Cohen's guidelines (Ellis, 2010), the obtained magnitude of the effects can be considered large.

Table 2
*Univariate descriptive statistics of the BIMS*

| Item | *M* | *SD* | *Skewness* | *Kurtosis* |
|---|---|---|---|---|
| **Behavior management (BM)** | | | | |
| 1. I nearly always intervene when students talk at inappropriate times during class | 5.20 | 0.96 | −1.29 | 1.56 |
| 2. I strongly limit student chatter in the classroom | 4.67 | 1.07 | −0.84 | 0.72 |
| 3. I am strict when it comes to student compliance in my classroom | 4.79 | 0.93 | −0.81 | 1.00 |
| 4. I firmly redirect students back to the topic when they get off task | 4.63 | 0.94 | −0.69 | 0.62 |
| 5. I insist that students in my classroom follow the rules at all times | 5.00 | 0.84 | −0.70 | 0.56 |
| 6. I strictly enforce classroom rules to control student behavior | 4.67 | 0.97 | −0.63 | 0.35 |
| 7. If a student's behavior is defiant, I will demand that they comply with my classroom rules | 4.80 | 0.95 | −0.93 | 1.09 |
| | | | | |
| **Instructional management (IM)** | | | | |
| 1. I nearly always use collaborative learning to explore questions in the classroom | 2.40 | 1.01 | 0.59 | 0.31 |
| 2. I engage students in active discussion about issues related to real world applications | 1.74 | 0.80 | 1.21 | 2.20 |
| 3. I nearly always use group work in my classroom | 3.25 | 1.15 | 0.28 | -0.22 |
| 4. I use student input when creating student projects | 2.59 | 1.10 | 0.71 | 0.54 |
| 5. I nearly always use inquiry-based learning in the classroom | 3.06 | 1.05 | 0.41 | 0.18 |
| 6. I nearly always adjust instruction in response to individual student needs | 2.45 | 0.93 | 0.47 | 0.49 |
| 7. I nearly always use a teaching approach that encourages interaction among students | 2.17 | 0.87 | 0.65 | 0.76 |

### Confirmatory Factor Analysis

First, a two-factor model with all 14 items was evaluated by calculating CFA. The goodness of fit indices suggested an ambiguous solution and a deviation of data from the theoretical model (Table 3). The original model of the scale (Martin & Sass, 2010) which does not include items BM3 and IM5 (model 2) also did not meet the desired standards. The third tested model was the one proposed in a validation study on a sample of teachers from Portugal (Sass et al., 2016). This model in which item BM2 and item IM6 were omitted yielded a good fit to the data based on CFI (.951), and an acceptable fit according to TLI (.940) and RMSEA (.076) values. After the analysis, the modification indices indicated that the item BM4 ("I firmly redirect students back to the topic when they get off task") cross-loaded on the IM factor. The modified two-factor model without this item was specified and tested. Although this modified model had the best fit indices (Table 3), as the model that included 12 items showed an acceptable fit to the data, the analysis of its psychometric characteristics continued.

Table 3
*Summary of the CFA and model fit indices BIMS*

| Model | Item | $\chi^2$ | df | CFI | TLI | RMSEA (90% CI) | WRMR |
|-------|------|---------|-----|------|------|------------------|-------|
| 1 | 14 [a] | 422.337 | 76 | .934 | .921 | .083 [.075–.091] | 1.506 |
| 2 | 12 [b] | 315.448 | 53 | .930 | .913 | .087 [.078–.096] | 1.433 |
| 3 | 12 [c] | 257.391 | 53 | .951 | .940 | .076 [.067–.086] | 1.293 |
| 4 | 11 [d] | 151.874 | 43 | .970 | .962 | .062 [.051–.073] | 1.046 |

*Note.* $\chi^2$ = chi-square; *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; 90% CI = 90% confidence interval for RMSEA; WRMR = weighted root mean square residual.

Models are listed in the following order:
[a] Model with fourteen items;
[b] Excluding items BM3 and IM5;
[c] Excluding items BM2 and IM6;
[d] Excluding items BM2, BM4, and IM6.

All standardized factor loadings were good and they were in the range from .49 to .86 (Table 4). In the current study 0.40 was accepted as the lowest factor loading limit (Tabachnick & Fidell, 2013). Correlation between the two factors was −.41.

Table 4
*Standardized factor loadings of the BIMS*

| Item | Standardized factor loading | Error |
|---|---|---|
| **Behavior management (BM)** | | |
| 1. I nearly always intervene when students talk at inappropriate times during class | .56 | .69 |
| 3. I am strict when it comes to student compliance in my classroom | .73 | .46 |
| 4. I firmly redirect students back to the topic when they get off task | .62 | .62 |
| 5. I insist that students in my classroom follow the rules at all times | .86 | .26 |
| 6. I strictly enforce classroom rules to control student behavior | .76 | .43 |
| 7. If a student's behavior is defiant, I will demand that they comply with my classroom rules | .64 | .59 |
| **Instructional management (IM)** | | |
| 1. I nearly always use collaborative learning to explore questions in the classroom | .67 | .55 |
| 2. I engage students in active discussion about issues related to real world applications | .53 | .72 |
| 3. I nearly always use group work in my classroom | .50 | .75 |
| 4. I use student input when creating student projects | .49 | .76 |
| 5. I nearly always use inquiry-based learning in the classroom | .60 | .64 |
| 7. I nearly always use a teaching approach that encourages interaction among students | .67 | .53 |

*Note.* All loadings are significant at $p < .001$ level.

### Reliability

Following the CFA, the reliability of the Serbian translation of the BIMS was analysed. The calculated Cronbach's alpha coefficients were .83 and .74 for BM and IM, respectively. Internal consistency was also analysed by calculating the omega coefficient for both subscales. The omega coefficient as an alternative measure of reliability has been recommended because Cronbach's alpha relies on the assumption of tau equivalence i.e. nearly identical factor loadings of all items on the scale (McNeish, 2018). The omega coefficients were: BM = .83, 95% CI [.81–.85] and IM = .74, 95% CI [.71–.77]. Generally speaking, the obtained results show that both subscales of the Serbian BIMS have adequate internal consistency.

### Convergent Validity

We examined the convergent validity by correlating the scores on the subscales of the BIMS with the dimensions of teachers' self-efficacy. Table 5 summarizes the results of the convergent validity analysis of the two-factor model of the BIMS. Correlations between IM and efficacy for student engagement, efficacy for classroom management and efficacy for instructional strategies were all significant and all in the expected direction. The BM scores were positively correlated with all three dimensions of teacher self-efficacy. Taken together, these findings supported the convergent validity of the BIMS.

Table 5
*Correlations of the two BIMS factors with the dimensions of teacher self-efficacy*

| Variable | Behavior management | Instructional management |
|---|---|---|
| Efficacy for student engagement | .24 | −.34 |
| Efficacy for classroom management | .34 | −.16 |
| Efficacy for instructional strategies | .21 | −.31 |

*Note.* All correlations are significant at $p < .01$ level.

## Group Differences in Teacher Behavior and Instructional Management

MANOVA was applied in order to examine the relationship between teacher gender, work experience, grade level and subscales of the BIMS. Significant differences were observed in the level of behavior and instructional management in relation to teachers' gender, Wilks lambda = .99, $F(2, 655)$ = 3.97, $p < .05$, partial $\eta^2$ = .012 (Table 6). A post hoc Bonferroni adjusted test showed that significant differences exist on both subscales of the BIMS, but in opposite directions ($p < .05$). Male teachers showed significantly higher levels of instructional management, while female teachers showed higher levels of behavior management.

There were no significant differences in dimensions of the BIMS in relation to the length of the teacher's work experience, Wilks lambda = .98, $F(8, 1308)$ = 1.73, $p > .05$, partial $\eta^2$ = .01.

Table 6
*Multivariate analysis of variance of the BIMS by the demographic variables of Serbian teachers*

| Demographic Variables | Behavior management | | | | | Instructional management | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *N* | *M* | *SD* | *F* | $\eta^2$ | *M* | *SD* | *F* | $\eta^2$ |
| Gender | | | | | | | | | |
| Male | 144 | 28.43 | 4.16 | 5.38 | .008 | 15.83 | 4.27 | 4.86 | .007 |
| Female | 514 | 29.30 | 3.91 | (1, 656)* | | 15.03 | 3.74 | (1, 656)* | |
| Teaching experience | | | | | | | | | |
| 1 year | 26 | 29.50 | 3.81 | | | 15.88 | 4.03 | | |
| 2–4 years | 47 | 29.51 | 3.59 | | | 15.57 | 4.31 | | |
| 5–9 years | 106 | 29.58 | 4.01 | 2.82 (4, 55) | .017 | 15.19 | 3.88 | 0.42 (4, 55) | .003 |
| 10–15 years | 125 | 29.78 | 3.43 | | | 14.97 | 4.10 | | |
| More than 15 years | 356 | 29.09 | 4.18 | | | 15.20 | 3.71 | | |
| Grade level | | | | | | | | | |
| Classroom teacher/ primary school | 173 | 29.66 | 3.88 | | | 14.34 | 3.35 | | |
| Subject teacher/ secondary school | 204 | 29.45 | 3.74 | 6.06 (2, 652)** | .018 | 15.56 | 3.50 | 6.07 (2, 652)** | .018 |
| High school teacher | 278 | 28.46 | 4.17 | | | 15.50 | 4.33 | | |

*Note.* $F$ = F test value; $\eta^2$ = partial eta squared.

* $p < .05$. ** $p < .01$.

MANOVA results showed that grade level has a significant effect on BIMS scores, Wilks lambda = .97, $F(4, 1302) = 5.46$, $p < .001$, partial $\eta^2 = .016$. As it is shown in Table 6, classroom teachers in elementary schools were found to score significantly higher on behavior management compared to those who work in high schools ($p < .01$). In addition, subject teachers who work in elementary schools showed significantly higher levels of behavior management compared to teachers who work in high schools ($p < .05$). The analysis showed that teachers who work in high schools report higher levels of instructional management in comparison with elementary school classroom teachers ($p < .01$). Furthermore, subject teachers in elementary school show higher levels of instructional management in comparison with classroom teachers. Bearing in mind the guidelines for the interpretation of the effect size in educational research (Ellis, 2010), the obtained values of partial $\eta^2$ indicated small effect sizes.

## Discussion

Numerous studies have shown that classroom management is an important determinant of students' academic, social and emotional learning (Evertson & Weinstein, 2006; Jennings & Greenberg, 2009). However, currently in Serbia there is a lack of standardized instruments that can be used in measuring teachers' approaches to management of student behavior and instruction. For these reasons, the purpose of this research was to assess the factor structure and the psychometric properties of the Serbian version of the BIMS.

Although the results from the CFA did not convincingly support the original two-factor model of the scale, a reasonably good fit with the data provided the model which was proposed in the study on a sample of Portuguese teachers (Sass et al., 2016). It appeared that the model which did not include the items BM2 ("I strongly limit student chatter in the classroom") and IM6 ("I nearly always adjust instruction in response to individual student needs") was better suited to Serbian teachers. The CFA showed that all items have significant standardized factor loadings (between .49 and .86), although it is noticeable that factor loadings on the IM subscale were somewhat lower. The results of this study suggested that classroom management could be adequately represented with two latent factors instead of one, which confirms the assumption of multidimensionality of the construct. However, the obtained negative correlation ($r = -.41$) of latent factors is not in line with previous research. In the initial validation study (Martin & Sass, 2010), the correlation between two factors was .22 which indicates that these two constructs are relatively independent. Additionally, the study on the sample of Portuguese teachers found correlations from .40 to .50, depending on the assessment method (Sass et al., 2016).

The obtained differences can be interpreted in multiple ways. First of all, it is possible that teachers in Serbia welcome constructivist methods that are based on enabling students to have active role in the process of acquiring knowledge and simultaneously tend to exhibit higher levels of control of student behavior. For example, the TALIS study (OECD, 2009) showed that in some countries

there is a positive correlation between constructivist and direct transmission approaches to teaching. Secondly, since teachers' work-related beliefs and instructive practices are under the strong influence of the national school system, culture and pedagogical traditions (Klassen et al., 2012), this negative correlation can be explained by a specific socio-cultural context. In particular, the findings of the present study suggest that for teachers in Serbia items on the BM subscale have a positive connotation while at the same time they reported that they use interactive teaching strategies. Contemporary literature suggests that students' active participation and effective behavior management are not mutually exclusive approaches and can be successfully combined (Kunter et al., 2013).

Reliability analysis showed that both subscales of the BIMS have adequate internal consistency. In this study, the alpha and omega estimates of the two factors were .83 (BM) and .74 (IM). These results suggest that Serbian researchers can reliably calculate and interpret the scores on both scales of the BIMS. In order to examine the convergent validity of the BIMS, relationships between BIMS subscales and teacher self-efficacy were examined. The convergent validity analysis showed low to moderate negative correlations between instructional management and efficacy for student engagement, efficacy for classroom management and efficacy for instructional strategies. This finding appears to be consistent with previous studies (Martin & Sass, 2010; Martin et al., 2012) that have shown that teachers who doubt their self-efficacy in this domain are more likely to use more controlling instructional strategies. Nie, Tan, Liau, Lau, & Chua (2013) reported that the relation between teacher efficacy and constructivist approach to instruction is stronger than the relation of teacher's self-efficacy and the traditional didactic approach. Therefore, our research confirms that high teacher self-efficacy contributes to student-centred, constructivist approaches to instruction.

Interestingly, behavior management had low to moderate positive correlations with all three components of teacher self-efficacy. Previous research on the relation of teachers' self-efficacy and behavior management has not led to consistent results. While there is an insight that behavior management is a predictor of teacher self-efficacy (Sass et al., 2016), some studies (Martin & Sass, 2010) have established a negative correlation between these two constructs. This discrepancy between the findings can be attributed to the characteristics of different cultural contexts (European samples vs. United States sample). It is quite likely that the construct of behavior control does not have the same meaning in different cultures. Therefore, its relations with other variables can be dependent on the national educational context.

We found that male teachers tend to exhibit more controlling approaches to instruction. These results further confirm previous research that has found male teachers to be more likely to directly control teaching while female teachers tended to use diversified instructional strategies (Lam et al., 2010). In this study, it has been found that female teachers exhibit higher levels of student behavior control. Results of previous studies (Hopf & Hatzichristou, 1999)

suggest that female teachers are more sensitive to external problems in student behavior, especially male adolescents, and that male teachers interpret student interpersonal behavior as less problematic in comparison to female teachers. Nonetheless, it should be noted that all differences in teachers' behavior and instructional management relating to gender had small effect sizes.

The results of the present study showed that elementary school classroom and subject teachers exhibit higher levels of behavior management than teachers who work in high schools. This can be interpreted by the developmental characteristics of high school students that have implications on their relationships with teachers. Developmental changes that occur in adolescence can create difficulties for teachers to control students' behavior (Ryan, Kuusinen, & Bedoya-Skoog, 2015). In addition, we found that subject teachers in secondary and high schools exhibit higher levels of instructional control in comparison with classroom teachers. We assume that these findings reflect the specifics of subject teaching in Serbia. Unlike class teachers, subject teachers spend less time with students in an environment that is performance oriented (Wang & Eccles, 2012) making it difficult for teachers to be committed to involving students in cognitively demanding activities based on the principles of active learning. Nonetheless, small effect sizes that were found in this study suggest that those differences in classroom management approaches based on grade level have low practical significance.

## Limitations and Future Research

One of the significant limitations of the current study is the inability to examine the relationships of the Serbian version of the BIMS with other relevant instruments that measure classroom management. Therefore, the future studies in the Serbian educational context should pay attention to the validation of other research tools that can be used in the measurement of classroom management and similar constructs. Further, it would be beneficial for the future research to examine the relationships of the BIMS with significant educational outcomes within the Serbian context, such as teacher stress, student motivation, and academic achievement.

## Conclusion

Based on the results of the present research, it can be concluded that the BIMS represents a valid and reliable instrument that can provide valuable information to researchers in Serbia about the classroom management of teachers. In addition to good psychometric characteristics, an important advantage of the BIMS is that it is a short scale that is easy to apply for scientific and practical purposes. The initial validation of the Serbian version of the BIMS provides opportunities to use this instrument in the educational context of Serbia and comparing results with findings obtained in other countries.

# References

Aloe, A. M., Amo, L. C., & Shanahan, M. E. (2014). Classroom management self-efficacy and burnout: A multivariate meta-analysis. *Educational Psychology Review*, *26*(1), 101–126. https://doi.org/10.1007/s10648–013–9244–0

Darling-Hammond, L., & Youngs, P. (2007). Defining "highly qualified teachers": What does "scientifically-based research" actually tell us? *Educational Researcher*, *31*(9), 13–25. https://doi.org/10.3102/0013189x031009013

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the weighted root mean square residual: evidence for trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 453–466. https://doi.org/10.1080/10705511.2017.1390394

Ellis, D. P. (2010). *The essential guide to effect size: Statistical power, meta-analysis, and the interpretation of research results*. New York City, NY: Cambridge University Press.

Evertson, C. M., & Weinstein, C. S. (2006). Classroom management as a field of inquiry. In C. S. Evertson, & C. M. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 3–15). Mahwah, NJ: Lawrence Erlbaum Associates.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York City, NY: Routledge.

Hopf, D., & Hatzichristou, C. (1999). Teacher gender-related influences in Greek schools. *British Journal of Educational Psychology*, *69*, 1–18.

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. https://doi.org/10.1080/10705519909540118

Jalali, S., Panahzade, V., & Firouzmand, A. (2014). Attitude towards computers and classroom management of language school teacher. *International Journal of Instruction*, *7*(2), 35–50. https://doi.org/10.1136/bmj.1.2632.1407-a

Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research*, *79*(1), 491–525. https://doi.org/10.3102/0034654308325693

Klassen, R. M., Aldhafri, S., Mansfield, C. F., Purwanto, E., Siu, A. F. Y., Wong, M. W., & Woods-Mcconney, A. (2012). Teachers engagement at work: An international validation study. *Journal of Experimental Education*, *80*(4), 317–337. https://doi.org/10.1080/00220973.2012.678409

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, *86*(3), 643–680. https://doi.org/10.3102/0034654315626799

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, *105*(3), 805–820. https://doi.org/10.1037/a0032583

Lam, Y. H., Tse, S. K., Lam, J. W. I., & Loh, E. K. Y. (2010). Does the gender of the teacher matter in the teaching of reading literacy? Teacher gender and pupil attainment in reading literacy in Hong Kong. *Teaching and Teacher Education*, *26*(4), 754–759. https://doi.org/10.1016/j.tate.2009.10.010

Martin, N. K., & Sass, D. A. (2010). Construct validation of the Behavior and Instructional Management Scale. *Teaching and Teacher Education*, *26*(5), 1124–1135. https://doi.org/10.1016/j.tate.2009.12.001

Martin, N. K., Sass, D. A., & Schmitt, T. A. (2012). Teacher efficacy in student engagement, instructional management, student stressors, and burnout: A theoretical model using in-class variables to predict teachers' intent-to-leave. *Teaching and Teacher Education*, *28*(4), 546–559. https://doi.org/10.1016/j.tate.2011.12.003

Martin, N. K., Yin, Z., & Mayall, H. (2006). *Classroom management training, teaching experience and gender: Do these variables impact teachers' attitudes and beliefs toward classroom management style?* Paper presented at the annual conference of the Southwest Educational Research Association, Austin, TX.

Martin, N. K., Yin, Z., & Mayall, H. (2008). The Attitudes and Beliefs on Classroom Control Inventory-revised and revisited: A continuation of construct validation. *Journal of Classroom Interaction*, *42*(2), 11–20.

McNeish, D. (2018). Thanks coefficient alpha, we' ll take it from here. *Psychological Methods*, *23*(3), 412–433.

Muthén, L. K., & Muthén, B. O. (2015). *Mplus statistical software version 7.31*. Los Angeles, CA: Muthén & Muthén.

Nie, Y., & Lau, S. (2009). Complementary roles of care and behavioral control in classroom management: The self-determination theory perspective. *Contemporary Educational Psychology*, *34*(3), 185–194. https://doi.org/10.1016/j.cedpsych.2009.03.001

Nie, Y., Tan, G. H., Liau, A. K., Lau, S., & Chua, B. L. (2013). The roles of teacher efficacy in instructional innovation: Its predictive relations to constructivist and didactic instruction. *Educational Research for Policy and Practice*, *12*(1), 67–77. https://doi.org/10.1007/s10671-012–9128-y

Ninković, S., & Knežević Florić, O. (2018). Validation of the Serbian version of the Teachers' Sense of Efficacy Scale (TSES). *Zbornik Instituta za pedagoška istraživanja*, *50*(1), 72–92. https://doi.org/10.2298/ZIPI1801072N

O'Neill, S. C., & Stephenson, J. (2011). The measurement of classroom management self-efficacy: A review of measurement instrument development and influences. *Educational Psychology*, *31*(3), 261–299. https://doi.org/10.1080/01443410.2010.545344

Oberle, E., & Schonert-Reichl, K. A. (2016). Stress contagion in the classroom? The link between classroom teacher burnout and morning cortisol in elementary school students. *Social Science and Medicine*, *159*, 30–37. https://doi.org/10.1016/j.socscimed.2016.04.031

OECD. (2009). *Creating effective teaching and learning environments: First results from TALIS*. Paris: OECD.

Peters, G. J. Y. (2014). The alpha and the omega of scale reliability and validity: Why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *European Health Psychologist*, *16*, 56–69. https://doi:10.31234/osf.io/h47fv

Ryan, A. M., Kuusinen, C. M., & Bedoya-Skoog, A. (2015). Managing peer relations: A dimension of teacher self-efficacy that varies between elementary and middle school teachers and is associated with observed classroom quality. *Contemporary Educational Psychology*, *41*, 147–156. https://doi.org/10.1016/j.cedpsych.2015.01.002

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment*, *29*(4), 347–363. https://doi.org/10.1177/0734282911406661

Sass, D. A., Lopes, J., Oliveira, C., & Martin, N. K. (2016). An evaluation of the Behavior and Instructional Management Scale's psychometric properties using Portuguese teachers. *Teaching and Teacher Education*, *55*, 279–290. https://doi.org/10.1016/j.tate.2016.01.020

Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education*, *62*(4), 339–355. https://doi.org/10.1177/0022487111404241

Tabachnick, B., & Fidell, L. (2013). *Using multivariate statistics* (6th ed.). Upper Saddle River, NJ: Pearson Education.

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805. https://doi.org/10.1016/S0742–051X(01)00036–1

Ünal, Z., & Ünal, A. (2012). The impact of years of teaching experience on the classroom management approaches of elementary school teachers. *International Journal of Instruction*, *5*(2), 1694–609.

van Buuren, S. (2018). *Flexible imputation of missing data*. New York City, NJ: CRC Press (Taylor & Francis Group).

Wang, M., & Eccles, J. S. (2012). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. *Child Development*, *83*(3), 877–895.

# Procena psihometrijskih karakteristika Skale upravljanja ponašanjem i nastavom: studija na uzorku srpskih nastavnika

## Stefan Ninković i Olivera Knežević Florić

*Odsek za pedagogiju, Filozofski fakultet, Univerzitet u Novom Sadu, Srbija*

Skala upravljanja ponašanjem i nastavom (eng. The Behavior and Instructional Management Scale – BIMS) je razvijena u cilju procene dvodimenzionalnog konstrukta upravljanja u učionici. Cilj ove studije je da se proveri faktorska struktura i psihometrijske karakteristike BIMS skale na uzorku srpskih nastavnika. Primenjena je konfirmativna faktorska analiza na uzorku od 660 nastavnika, a rezultati podržavaju dvofaktorski model latentne strukture skale. Obe subskale su imale zadovoljavajuću internu konzistentnost. Rezultati su pokazali i dobru konvergentnu validnost dvofaktorskog modela BIMS skale. Zaključujemo da se upotreba BIMS skale može preporučiti istraživačima zainteresovanim za merenje upravljanja u učionici nastavnika u Srbiji.

**Ključne reči**: upravljanje ponašanjem, upravljanje nastavom, nastavnici, upravljanje u učionici, konfirmativna fakotorska analiza, Srbija