

Martinez-Gomez, Monica; Marin-Garcia, Juan A.; Girado-O'Meara, Martha

## Article

# The measurement invariance of job diagnostic survey (JDS) across three university student groups

Journal of Industrial Engineering and Management (JIEM)

## Provided in Cooperation with:

The School of Industrial, Aerospace and Audiovisual Engineering of Terrassa (ESEIAAT), Universitat Politècnica de Catalunya (UPC)

*Suggested Citation:* Martinez-Gomez, Monica; Marin-Garcia, Juan A.; Girado-O'Meara, Martha (2016) : The measurement invariance of job diagnostic survey (JDS) across three university student groups, Journal of Industrial Engineering and Management (JIEM), ISSN 2013-0953, OmniaScience, Barcelona, Vol. 9, Iss. 1, pp. 17-34, <https://doi.org/10.3926/jiem.1783>

This Version is available at:

<https://hdl.handle.net/10419/188760>

## Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

## Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc/3.0/>

## The Measurement Invariance of Job Diagnostic Survey (JDS) Across Three University Student Groups

Mónica Martínez-Gómez<sup>1</sup>, Juan A. Marin-García<sup>1</sup>, Martha Girado-O'Meara<sup>2</sup>

<sup>1</sup>Universitat Politècnica de Valencia, <sup>2</sup>Facultad de Psicología, Universidad de Valencia (Spain)

[momargo@eio.upv.es](mailto:momargo@eio.upv.es), [jamarin@omp.upv.es](mailto:jamarin@omp.upv.es), [marthaomeara@gmail.com](mailto:marthaomeara@gmail.com)

Received: November 2015

Accepted: January 2016

### **Abstract:**

**Purpose:** The main purpose of this study is to apply a multigroup confirmatory analysis to examine the measurement invariance (MI) of the adapted version of the Job Diagnosis Survey (JDS) as a measurement tool that analyses the relationship between the features of teaching methodologies with university students' motivation and satisfaction across data collected on different degrees and academic years.

**Design/methodology/approach:** Confirmatory factor analysis was carried out using a multigroup structural equation model, using the program EQS 6.1 to test the invariance of the adapted version of JDS in a sample constituted by 535 student of a Spanish public university. The assessment of invariance included the levels of configural, metric, scalar, covariance and latent variables invariance. Several goodness-of-fit measures were assessed.

**Findings:** The results show that measurements are equivalent at the configural, metric, covariance and latent factors invariance. Although the hypotheses of scalar invariance is rejected, results suggest that JDS is partial strict invariant and has satisfactory psychometric properties on all samples.

**Research limitations/implications:** The sample is framed in university students aged between 18 and 30 and for a questionnaire on teaching methodology and students' satisfaction

in the context of a Spanish university and the generalization to other questionnaire, or population, should be proved with specific data. Furthermore, the sample size is rather small.

**Originality/value:** In the current process of change that is taking place in universities according to the plan developed by the European Space of Higher Education, focused on increasing the student skills, validate instruments as the satisfaction scale of JDS, are necessary to evaluate students' satisfaction with new active methodologies. These findings are useful for researchers since they add the first sample in which the MI of a student's satisfaction survey is tested.

**Keywords:** measurement equivalence, student's satisfaction and motivation, Job Diagnosis Survey, multigroup confirmatory analysis, higher education, invariance

---

## 1. Introduction

Student satisfaction is a concept that has become more prevalent in higher education, since it is related to motivation, and to learning outcomes (Richardson, 2005). Satisfied students tend to work harder in their educational activities (Tessema, Ready & Yu, 2012), to perform better in these activities (Cotton, Dollard & Jonge, 2002), and to finish the academic year (Mason, 2012).

Satisfaction has been widely investigated, both in academic and enterprise world (Alves & Raposo, 2009; Van Saane, Sluiter, Verbeek & Frings-Dresen, 2003). It is a very complex concept and, despite all the research around it, in education, there is not a clear definition in order to measure it (García-Aracil, 2009; Tessema et al., 2012).

In the last decades, it has been an increased interest about how to enlarge the level of participation of the students in the process of learning. For that purpose, is necessary to change traditional teaching towards a more active educational methodology where the characteristics of the context of student learning are taken into account and be designed so, that increases motivation and student satisfaction. Many universities have incorporated some instruments to measure student satisfaction, however, little attention has been given to the structure and psychometric properties of this satisfaction surveys. Only few studies have validated students satisfaction surveys and none have evaluated the comparability of the questionnaire across other cultures, languages or universities.

The main purpose of this study is to apply a multigroup confirmatory analysis to examine the measurement invariance (MI) of the Job Diagnostic Survey (JDS) adapted to teaching, which incorporates items about teaching methodology and items about student's motivation and satisfaction, across data collected on different degrees and academic years.

This study is structured as follows. First, the theoretical framework of how to measure student satisfaction. Second, the background of measurement equivalence (MI) of satisfaction students' surveys across different groups. Third, we present the research methodology and the results obtained. Finally, this paper concludes with the main reflection of findings achieved in our analysis, their limitations and recommendations for further research.

## **2. Theoretical Framework**

### **2.1. How to Measure Student Motivation and Satisfaction**

There are many and varied instruments to measure student satisfaction, however, Beltyukova and Fox (2002) noted that all satisfaction instruments use different satisfaction subscales determined by different items. Beltyukova and Fox (2002) further explained that there are different instruments to evaluate student satisfaction but each one is related to a different definition of this concept which is a frequent criticism on the researches. From a learning perspective, satisfaction is related to student perceptions about characteristics of academic activities, autonomy (Mason, 2012), relationship with the teacher, degree of customization of the class, meeting goals and teamwork (Reinig, Horwitz & Whittenburg, 2011). Moreover, several studies have investigated the factors related to student motivation and the influence in their satisfaction (Adler, Milne & Stablein, 2001; Ames, 1992; Paris & Turner, 1994; Ünal & Inan, 2010).

Over the last few years, reflections on the analogy between the business and academic worlds has been studied in various disciplines (i.e. Adler et al., 2001; Armstrong, 2003; Cotton et al., 2002; DeShields, Kara & Kaynak, 2005; Donaldson, 2002; Freed, 2005; Martínez-Gómez & Marín-García, 2009; O'Neil & Hopkins, 2002), especially in business management and engineering.

The identification of the need for improvement the job, working so that increase satisfaction and motivation of employees is called in the business world redesign of jobs. The purpose is to create a more motivating job by increasing the level of autonomy, significance, variety, feedback, identity with the product or service carried out, and social contact (Hackman & Oldham, 1980). One of the tools used to measure and guide the process of redesigning job positions, is the model of the job characteristics of Hackman and Oldham (1975) and the questionnaire derived from this model, the Job Diagnosis Survey (JDS).

In the analogy between the business and academic worlds, we consider the teacher as a leader, managing a group of people (French, 2006) and we can consider our subjects as employees. The job designing is related to the teaching methodology and we can use tools that are useful in business management to measure students motivation and satisfaction. In this

paper, we used the academic version of the global Job Diagnostic Survey (JDS), a validated questionnaire (Martínez-Gómez & Marín-García, 2009), as a diagnosis tool of active methodologies to measure students' satisfaction.

## **2.2. Research on Measurement Invariance on Teaching's Students Surveys**

A major task ahead is to establish whether the instruments to measure students' satisfaction are invariant across different situations and with different groups of people. Jöreskog (1971) was the first to question the need to test equivalence between factor structures and recommended that all equivalence tests began with a global analysis of the covariance structures across groups. Byrne (1989) established the importance of assessing if means across different groups were equivalent, but researchers seldom tested it. In the same line, Elosua (2005) argues that the comparison between findings obtained with a test in different contexts, can only be accepted if their formal and substantive equivalence is empirically justified.

A review of efforts to assess the invariance of measurement instruments across different respondent groups using confirmatory factor analysis is provided for the years (Vandenberg & Lance, 2000). Afterwards, Schmitt and Kuljanin (2008) developed a review to identify articles that used CFA in assessing factor invariance but the issues addressed in those studies were quite diverse.

Nowadays, the most common way to assess MI is by a multigroup confirmatory factor analysis, an aspect highly studied by Cheung and Rensvold (2002) and subsequently widened by Cheung (2008) for second-order models.

Few formal studies related to MI of satisfaction's questionnaires have been conducted at the moment and many of them are for test MI across different data collection modes. Leung and Kember (2005), carried out a study of the program evaluation questionnaire from students of Hong Kong university, in order to assess if the mode of administering the questionnaire had an effect in the interpretation of the items. They concluded that there are not different between both modes, with same scales reliability across samples, with same common factor structure and, finally, that factor structure of the model is invariant, except in the case of factors variance invariance. Later, Peinado, Chávez, Viciano and Rivero (2012) studied the psychometric equivalence of the Self-Efficacy scale in Academic Behaviour (EACA) in health sciences and social sciences students, revealing that metric invariance and strong factorial invariance (across intercepts) was met. Klieger, Centra, Young, Holtzman and Kotloff (2014) developed a review about students evaluation teachers surveys (SET) and the comparability of the online and paper versions using the SIR II™ Student Instructional Report (SIR II) to examine interrater reliability between both surveys modes.

Finally, other studies use samples of students to test the invariance across cross-cultures or cross-nations. For example, Li, Saklofske, Bowden, Yan and Fung (2012), assessed the measurement equivalence of the Wong & Law Emotional Intelligence Scale (WLEIS) with three groups of Chinese university students.

We have been unable to locate any studies that have investigated MI of teaching evaluation questionnaires across different career programmes. So, there is a lack of empirical studies that have addressed MI issues concerning students' satisfaction across multiple contexts.

### **2.3. Purpose and Contributions of Present Study**

The purpose of this study is to analyse MI of academic version of JDS when is completed by students of different degrees and academic years. This study represents the first attempt to investigate the MI of a student satisfaction survey. A progressive complete invariance was developed, so the researchers can find the cause of failure of the structural parameter considered in each stage (Elosua, 2005).

## **3. Methodology**

### **3.1. Sample**

The total sample was constituted by 535 student of a Spanish public university. The first group was from Faculty of Business Administration and Management (FBAM) from academic year 2008-2009 (N=205). The second group, was also from Faculty of Business Administration and Management but from the academic year 2009-2010 (N=189). Finally, the third group was from Faculty of Industrial Engineering (FIE) from academic year 2007-2008 (N=141). Dates of all samples were from different subjects and different courses. Questionnaires were completed as well in paper as web surveys. In a prior studied we established MI between both modes of administering the questionnaire, so we can use them jointly.

### **3.2. Instrument**

We selected the validated scale of the Job Diagnostic Survey (JDS) in its academic version (Martínez-Gómez, Marín-García & Girado-O'Meara, 2013) to examine MI across students of different degrees and academic years. The JDS (Hackman & Oldham, 1975) is one of the main tools to evaluate how stimulating a job position is (Fried, 1991; Griffin, 1991; De Treville & Antonakis, 2006). Its adapted version is a tool to diagnose teaching methodologies in

university and includes a satisfaction single-item scale, the Motivating Potential Score (MPS) and job characteristics scales (Figure 1).

Satisfaction with the students' workplace is measured by a single item (Overall, I am very satisfied with the course), with 7 levels of response, where higher values indicate greater satisfaction with the course.

The job characteristics are measured by six scales assessing the variety, identity, significance, autonomy, feedback from the job itself and feedback from agents (Figure 1). The scales contain 14 items measured in a seven-points Likert scale. The response scale is presented in two ways: one, measured with three items with anchorage phrases in a graphic scale (items of Section 1, S1). The second one, a phrase that they have to answer according to the level of agreement or disagreement (items of section 2, S2). The sum of the 6 characteristics of the job is the motivating potential score (MPS).

Each of the items of the features of the teaching method is valued on a Likert scale with 7 levels of response (very inappropriate - very appropriate).

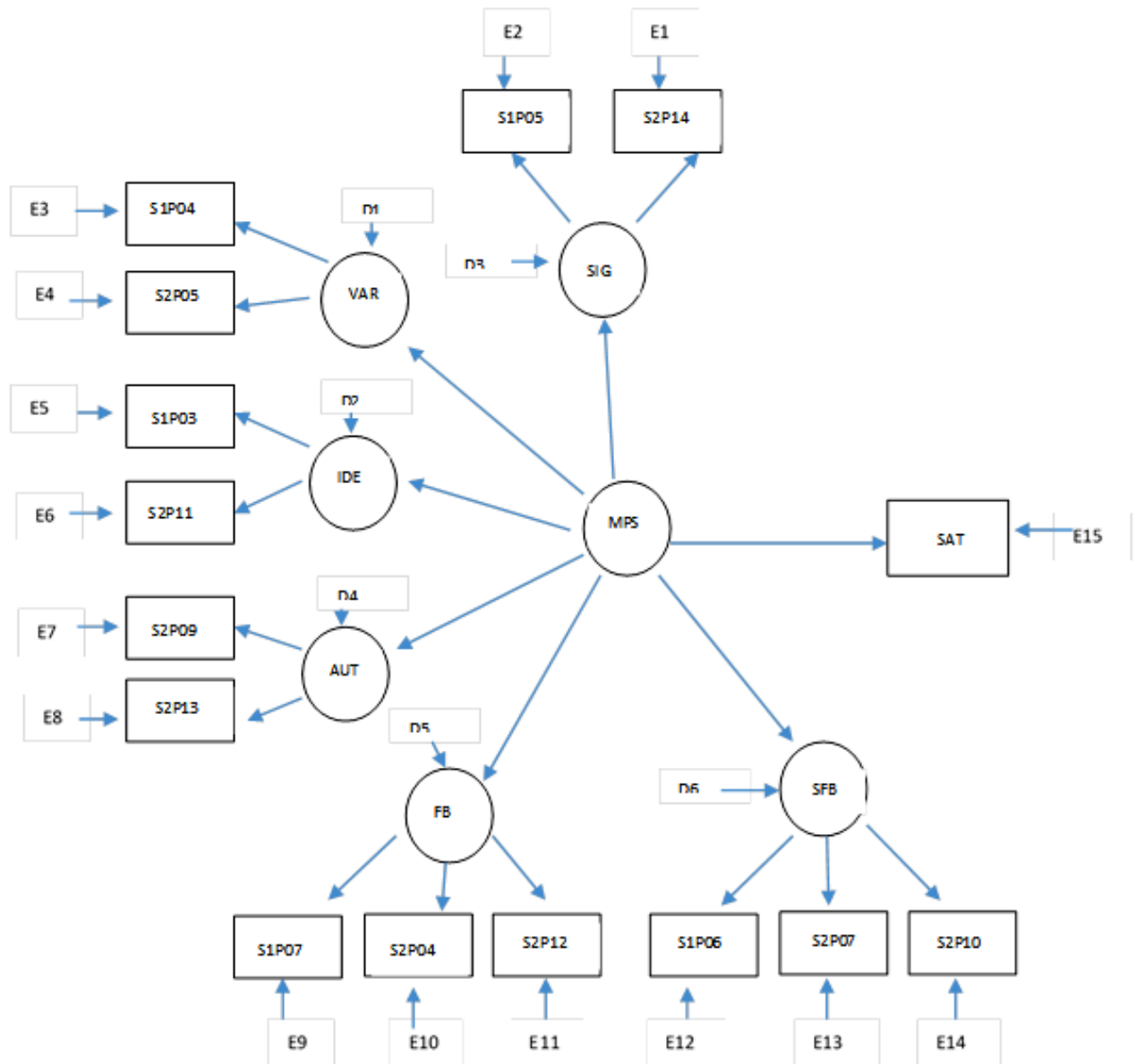


Figure 1. Model for the second-order factor model of teaching adapted version of JDS (Latent factor with acronym in parentheses. "significance" (SIG); "variety" (VAR); IDE=identity; AUT=autonomy; FB=feedback from the job itself; SFB=feedback from agents; SAT=Satisfaction. Items or observable variable were reverse coded are denoted with SiPi)

### 3.3. Method of Analysis

The first step in testing measurement invariance is to assess whether the factor structure of the academic version of JDS can be replicated across different groups. A confirmatory factor analysis (CFA) examined whether the second-order factor model of teaching adapted version of JDS validated in previous studies, (Martínez-Gómez et al., 2013) with the total sample, can be replicated in each of the three groups. Since the sample size is not very large and the  $\chi^2$  is very sensible with respect to sample variations (Bentler & Bonett, 1980; Bollen & Long, 1993; Browne & Cudek, 1993; Santos-Rego, Godás-Otero, Lorenzo-Moledo & Gómez Fraguera, 2010),



we have decided to include other additional indexes which compare discrepancies between the baseline model and the proposed one. The goodness of fit index (GFI), the adjusted goodness of fit index (AGFI) and the McDonald fit index (MFI) which show the relative amount of variance and covariance, comparing the observed sample matrix and the reproduced one. For a good fit, the GFI and the AGFI has to be above .85, the same as the comparative CFI. The incremental fit index (IFI), the Bentler and Bonnet's (1980) normed fit index and the non-normed fit index (NNFI) by the same authors, compare the fit function of the null model with the proposed model. For a good fit, they have to be above .90. The root mean square error of approximation (RMSEA) measures the discrepancy between expected and observed covariance matrix and values have to be below .08 (Ullman & Bentler, 2004).

The next step is to assess MI across groups. The invariance testing process involves several steps of increasingly restrictive measurement assumptions to evaluate the significance of the decrement of these indexes for each of the more constrained model (Byrne & Stewart, 2006; French & Finch, 2008; Millsap & Yun-Tien, 2004; Vandenberg & Lance, 2000; Widaman & Reise, 1997). We applied the multistep process suggested by Steenkamp and Buamgarther (1998).

We employed multigroup confirmatory factor analysis (MGCFA) using the programme EQS 6.1 and robust maximum likelihood estimation. As goodness of fit indexes, we used the Sattorra-Bentler scaled chi-square adjusted to nonnormality ( $SB\chi^2$ ) with robust standard errors (Sattorra & Bentler, 1994, 2001). The  $SB\chi^2$  works appropriately with all type of sample sizes and several researchers have recommended using it for multivariate nonnormal data (Curran, West & Finch, 1996). Although the  $SB\chi^2$  difference statistic has been the primary criterion, we used other index in this study, such as, the Robust Comparative Fit Index (RCFI) whose value should be above .90 for a good fit. Also, following Hu and Bentler (1999) considerations, we used the RMSEA (Standardized Root Mean Square Residual), whose value should be less than .08 to indicate a good fit (Hair, Anderson, Thatam & Black, 1998). Nevertheless, some authors (i.e. Cheung & Rensvold, 2002; Byrne & Stewart, 2006; Chen, 2007) argued that it is still possible to use these fit indices to test for measurement equivalence, but focusing on the changes in these measures when adding the constraints at the different steps. They consider that a change in CFI than .01 or less indicate that the invariance hypothesis should not be reject and Vandenberg and Lance (2000) suggest that when comparing successive invariance models across groups, a change in RMSEA values below .06, reflect excellent fit.

## 4. Results and Discussion

### 4.1. Confirmatory Factor Analysis

The results of CFA revealed that the second-order factor model was replicated in the three groups (see Table 2). Although the value of  $\chi^2$  is non-significant in all cases (p-value=.00021, .00000, .00126, .00000) the values of the other indexes are above or very close to the limit value for a good fit, except in the sample of FBAM 0809, where MFI and RMSEA are lower and upper its ideal values (MFI=.714 and RMSEA=.090).

Goodness of fit index	FBAM 08-09	FBAM 09-10	FIE	Total Sample
$\chi^2$	137,457	189,704	128,604	239,578
Df	84	84	84	84
p-value	.00021	.00000	.00126	0,00
GFI	.881	.869	.893	.929
AGFI	.830	.810	.847	.898
MFI	.822	.714	.853	.836
NFI	.869	.820	.798	.907
NNFI	.929	.861	.895	.921
CFI	.943	.889	.916	.937
RMSEA <sup>1</sup>	.069	.090	.062	.065

Table 1. Goodness of fit indexes of the three subgroups and the total sample

### 4.2. Configural Invariance

We tested configural invariance across surveys modes. We began by equality of means, to continue with equality of variances and covariances matrixes. As Table 3 shows, the value of  $SB\chi^2$  (p-value=.0000) for means and variances covariances don't support the equally assumption. In such cases, Satorra proposes to study the invariance of both parameters jointly. These results are shown in Table 4. Although the value of  $SB\chi^2$  (p-value=.00726) does not allow the hypothesis of invariance, the rest of indexes contradict this conclusion. The Robust configural fit index (RCFI=.955) allow us, with caution, accept the equality of the number of factors and factor pattern matrices.

Model	$\chi^2$ (p-value)	$SB\chi^2$ (p-value)	Df	RMSEA	CFI	Robust RMSEA	Robust CFI
Equality of means	2925,431 (.00000)	2493,328 (.00000)	345	.232	n.a	.212	n.a
Equality of covariances and variances	138,209 (.00000)	118,1128 (.00000)	31	.155	.956	.140	.953

Table 2. Goodness of fit indexes for configural invariance (means and variances-covariances)

Model	$\chi^2$ (p-value)	SB $\chi^2$ (p-value)	Df	RMSEA	CFI	Robust RMSEA	Robust CFI
Equality of means and variances	226,122 (.00282)	210,4562 (.00726)	31	.163	.955	.131	.950

Table 3. Goodness of fit indexes for invariance of means and variances together

### 4.3. Metric Invariance

As configural invariance can be established, we evaluated metric invariance across surveys models, constraining factor loadings in each group separately. As shown in Tables 5 and 6, the value of SB $\chi^2$  change (p-value=.041850757) is significant with a 90% confidence level and allow us to accept that the nested model was still well-fitting. Therefore we could not reject the hypothesis null.

Model	$\chi^2$ (p-value)	SB $\chi^2$ (p-value)	Df	RMSEA	CFI	Robust RMSEA	Robust CFI
Metric Invariance	374,736 (.00000)	319,6106 (.00005)	218	.071	.935	.057	.946
Metric Invariance without constraints	354,286 (.0000)	295,0138 (.00000)	204	.072	.93	.056	.951

Table 4. Goodness of fit indexes for metric

Satorra-Bentler Scaled Difference	D.f.	p-value
24,3246	14	.041850757

Table 5. Difference of adjusted Satorra-Bentler Chi Squared indexes

### 4.4. Scalar Invariance

Next, we evaluated if scalar invariance can be established constraining the intercepts of all surveys modes. As shown in Tables 7 and 8, the value of SB $\chi^2$  change is very significant (p<.001), which indicates that the scalar invariances was not supported. However, if we consider again the value of Robust CFI (RCFI=.939) and Robust RMESA (.061), we can cautiously establish that there is scalar invariance across the three groups.

Model	$\chi^2$ (p-value)	SB $\chi^2$ (p-value)	Df	RMSEA	CFI	Robust RMSEA	Robust CFI
Scalar Invariance	497,411 (.0000)	396,2519 (.0000)	222	.092	.900	.073	.924
Scalar Invariance without constraints	470,119 (.00000)	326,5043 (.00001)	192	.092	.894	.061	.939

Table 6. Goodness of fit indexes for scalar invariance

Satorra-Bentler Scaled Difference	D.f.	p-value
69,7476	30	5,24741E-11

Table 7. Difference of adjusted Satorra-Bentler Chi Squared indexes

#### 4.5. Covariance Invariance Among Latent Factors

The next step is to test if there are differences in covariance among latent factors across the three groups. Since scalar invariance could be established with caution, we conducted this test, imposed restrictions on the metric invariance model. Table 9 and 10 show the results. This comparison yielded a value of SB $\chi^2$  change significant, (p-value=.12299) and we can state that there is covariance invariance across groups.

Model	$\chi^2$ (p-value)	SB $\chi^2$ (p-value)	Df	RMSEA	CFI	Robust RMSEA	Robust CFI
Covariance Invariance	392,74 (.00000)	332,3300 (.00000)	226	.072	.931	.057	.943
Metric Invariance	374,737 (.00000)	319,6106 (.00001)	218	.071	.935	.057	.963

Table 8. Goodness of fit indexes for covariance invariance

Satorra-Bentler Scaled Difference	D.f.	p-value
12,6894	8	.122992637

Table 9. Difference of adjusted Satorra-Bentler Chi Squared indexes

#### 4.6. Variance Invariance Across Latent Factors

To evaluate variance invariance of latent factors is necessary adds a new restriction about the standard errors across survey modes. If we can establish factor latent variance invariance across groups, as covariance invariance have yet established, correlation across latent factors will be the same across groups, which means that the relation of the factors with the MPS is

the same in the original model, independent of the group. Results are showed in Tables 11 and 12. As the change of p-value achieved when comparing the  $SB\chi^2$  index is .01556, we can accept that invariance across latent factors is equivalent across samples with 90% confidence level, as well as the values of the robust CFI (RCFI=.936) and robust RMSEA (RRMESA=.055).

Model	$\chi^2$ (p-value)	$SB\chi^2$ (p-value)	Df	RMSEA	CFI	Robust RMSEA	Robust CFI
Invariance factors variance	392,74 (.00000)	332,3300 (.00000)	226	.072	.931	.057	.943
Invariance errors variance	417,174 (.00000)	36,0362 (.00000)	240	.072	.927	.059	.936

Table 10. Goodness of fit indexes for latent factors variance invariance

Satorra-Bentler Scaled Difference	D.f.	p-value
27,706	14	.015561284

Table 11. Difference of adjusted Satorra-Bentler Chi Squared indexes

#### 4.7. Variance Invariance of Errors

Finally, we analyzed the invariance of measurement errors variance across samples. In this case, as we can see in Table 13 and 14, the p-value for the change of  $SB\chi^2$  is .66187, so we can state that reliability of the surveys items is similar across the students of the three degree programmes.

Model	$\chi^2$ (p-value)	$SB\chi^2$ (p-value)	Df	RMSEA	CFI	Robust RMSEA	Robust CFI
Covariance Invariance	417,174 (.00000)	360,0362 (.00001)	240	.072	.927	.059	.936
Latent factors variance invariance	447,890 (.00000)	384,4060 (.00000)	268	.068	.926	.055	.939

Table 12. Goodness of fit indexes for errors variance invariance of latent factors

Satorra-Bentler Scaled Difference	D.f.	p-value
24,3698	28	.661874583

Table 13. Difference of adjusted Satorra-Bentler Chi Squared indexes

The results indicate that we can state that JDS adapted to university teaching reveal the same factor structure, factor loadings, factors latent variances and reliability of scales. This mean

that we can integrate data collected from different degrees programmes to increase the response rate, following the recommendation of the Standards for Educational Research Association (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1999).

## **5. Conclusions**

The purpose of our study was to evaluate if the underlying factor structure of the teaching version of JDS was equivalent with data collected across different samples. Results show that the global model proposed to evaluate the relationship between satisfaction, MPS and job characteristics is invariant across the samples, regarding to configural, metric, covariance, latent factors and measurement error of latent factors invariance. Chen (2007) established there would be not such a big difference across the groups, if the scale would have offered a similar reliability in all groups. So, the non-fulfilment of scalar invariance is not important for the purpose of this study.

For these reasons, we can conclude that the three samples can be considered equivalent, with the same factor structure, factor loadings, measurement errors of factors and the same reliability even when complete invariance could not be proved. As Van de Schoot, Lugtig and Hox (2012) set out, when checking if factor loadings, items coefficients and residual variances are equivalent across groups, we can state that comparisons made across groups are valid at all levels.

These findings are useful for researchers since they add the first sample in which the MI of a student's satisfaction survey is tested. This sample is framed in university students aged between 18 and 30 and for a questionnaire on teaching methodology and students' satisfaction in the context of a Spanish university. We also offer the procedure to follow to test MI with progressive complete invariance test.

There were of course, limitations to this study. As stated previously, we used a student sample with a specific questionnaire and the generalization to other questionnaire, or population, should be proved with specific data. Furthermore, the sample size is rather small. It is possible that analyses based on larger sizes or different context or universities would yield different results. That determination requires additional investigation.

Finally, we only tested measurement invariance of a second-order model but we should tested invariance in first-order models since although second-order models fit well against invariance, the theoretical explanation of MI is always ambiguous.

As future research lines, first we extend the sample to a representative population of the university students' population (Spanish or other countries). Secondly, we analyse first-order models invariance.

## References

- Adler, R.W., Milne, M.J., & Stablein, R. (2001). Situated Motivation: An Empirical Test in an Accounting Course. *Canadian Journal of Administrative Sciences*, 18, 101-115. <http://dx.doi.org/10.1111/j.1936-4490.2001.tb00248.x>
- Alves, H., & Raposo, M. (2009). The measurement of the construct satisfaction in higher education. *The service industrial journal*, 29, 203-218. <http://dx.doi.org/10.1080/02642060802294995>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Ames, C. (1992). Classrooms: Goals, Structures, and Student Motivation. *Journal of Educational Psychology*, 84, 261-271. <http://dx.doi.org/10.1037/0022-0663.84.3.261>
- Armstrong, M.J. (2003). Students as Clients: A Professional Services Model for Business Education. *Academy of Management Learning and Education*, 2, 371-374. <http://dx.doi.org/10.5465/AMLE.2003.11901964>
- Beltyukova, S.A., & Fox, C. (2002). Student satisfaction as a measure of student development: Towards a universal metric. *Journal of College Student Development*, 43, 161-172.
- Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606. <http://dx.doi.org/10.1037/0033-2909.88.3.588>
- Bollen, K.A., & Long, J.S. (Eds.). (1993). *Testing structural equation models*. Newbury Park, California: Sage.
- Brown, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In Bollen, K.A., & Long, J.S. (Eds.). *Testing structural equation models*. Newbury Park, California: Sage. 136-162.
- Byrne, B.M. (1989). Multigroup comparisons and the assumption of equivalent construct validity across groups: Methodological and substantive issues. *Multivariate Behavioural Research*, 24, 503-523. <http://dx.doi.org/10.1037/0033-2909.88.3.588>
- Byrne, B.M., & Stewart, S.M. (2006). The MACS approach to testing for multigroup invariance of a second-order structure: a walk through the process. *Structural Equation Modeling*, 13(2), 287-321. [http://dx.doi.org/10.1207/s15328007sem1302\\_7](http://dx.doi.org/10.1207/s15328007sem1302_7)
- Chen, F.F. (2007). Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling*, 14, 464-504. <http://dx.doi.org/10.1080/10705510701301834>

- Cheung, G.W., & Rensvold, R.B. (2002). Evaluating goodness-of-fit indices for testing measurement Equivalence. *Structural Equation Modeling*, 9, 233-255. [http://dx.doi.org/10.1207/S15328007SEM0902\\_5](http://dx.doi.org/10.1207/S15328007SEM0902_5)
- Cheung, G.W. (2008). Testing equivalence in the structure, means, and variances of higher-order constructs with structural equation modeling. *Organizational Research Methods*, 11(3), 593-613. <http://dx.doi.org/10.1177/1094428106298973>
- Cotton, S., Dollard, M., & de Jonge, J. (2002). Stress and Student Job Design: Satisfaction, Well-Being, and Performance in University Students. *International Journal of Stress Management*, 9, 147-162. <http://dx.doi.org/10.1023/A:1015515714410>
- Curran, P.J.; West, S.G., & Finch, J.F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(11), 16-29. <http://dx.doi.org/10.1037/1082-989X.1.1.16>
- DeShields, O.W., Kara, A., & Kaynak, E. (2005). Determinants of business student satisfaction and retention in higher education: Applying Herzberg's two-factor theory. *International*, 19, 128-139. <http://dx.doi.org/10.1108/09513540510582426>
- De Treville, S., & Antonakis, J. (2006). Could Lean production job design be intrinsically motivating? Contextual, configurational, and levels-of-analysis issues. *Journal of Operations Management*, 24, 99-123. <http://dx.doi.org/10.1016/j.jom.2005.04.001>
- Donaldson, L. (2002). Damn by Our Own Theories: Contradictions Between Theories and Management Education. *Academy of Management Learning and Education*, 1, 96-106. <http://dx.doi.org/10.5465/AMLE.2002.7373701>
- Elosua, P. (2005). Evaluación progresiva de la invarianza factorial entre las versiones original y adaptada de una escala de autoconcepto. *Psicothema*, 17(2), 356-362.
- Freed, J.E. (2005). Creating a Total Quality Environment (TQE) for Learning. *Journal of Management Education*, 29, 60-81. <http://dx.doi.org/10.1177/1052562904264283>
- French, B.F., & Finch, W.H. (2008). Multigroup confirmatory factor analysis: Locating the invariant referent sets. *Structural Equation Model*, 15(1), 96-113. <http://dx.doi.org/10.1080/10705510701758349>
- French, N.K., & Chopra, R.V. (2006). Teachers as executives. *Theory into Practice*, 45, 230-238. [http://dx.doi.org/10.1207/s15430421tip4503\\_5](http://dx.doi.org/10.1207/s15430421tip4503_5)
- Fried, Y. (1991). Meta-analytic comparison of the job diagnostic survey and job characteristics inventory as correlates of work satisfaction and performance. *Journal of Applied Psychology*, 76(5), 690-697. <http://dx.doi.org/10.1037/0021-9010.76.5.690>



- García-Aracil, A. (2009). European graduates' level of satisfaction with higher education. *Higher Education*, 57, 1-21. <http://dx.doi.org/10.1007/s10734-008-9121-9>
- Griffin, R.W. (1991). Effects of work redesign on employee perceptions, attitudes and behaviours: A long-term investigation. *Academy of Management Journal*, 34, 425-435. <http://dx.doi.org/10.2307/256449>
- Hackman, J.R., & Oldham, G.R. (1975). Development of the Job Diagnostic Survey. *Journal of Applied Psychology*, 60(2), 159-170. <http://dx.doi.org/10.1037/h0076546>
- Hackman, J.R., & Oldham, G.R. (1980). *Work Redesign*. Addison-Wesley.
- Hair, J.F., Anderson, R.E., Thatam, R.L., & Black, W.C. (1998) *Multivariate data analysis*. 6th edition. New York, Prentice Hall International.
- 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. <http://dx.doi.org/10.1080/10705519909540118>
- Jöreskog, K.G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36, 409-426. <http://dx.doi.org/10.1007/BF02291366>
- Klieger, D., Centra, J., Young, J., Holtzman, S., & Kotloff, J.L. (2014). *Testing the invariance of interrater reliability between paper-based and online modalities of the SIR II™ student instructional report*. Available at: <http://www.ets.org/Media/Research/pdf/SIRII-Report-Klieger-Centra-2014.pdf>
- Leung, D., & Kember, D. (2005). Comparability of data gathered from evaluation questionnaires on paper and through the internet. *Research in Higher Education*, 46(5), 571-591. <http://dx.doi.org/10.1007/s11162-005-3365-3>
- Li, T., Saklofske, D.H., Bowden, S.C., Yan, G., & Fung, T.S. (2012). The measurement invariance of the Wong and Law Emotional Intelligence Scale (WLEIS) across three Chinese university student groups from Canada and China. *Journal of Psychoeducational Assessment*, 30(4), 439-452. <http://dx.doi.org/10.1177/0734282912449449>
- Martínez-Gómez, M., & Marín-García, J.A. (2009). Como medir y guiar el cambio hacia entornos educativos universitarios más motivadores para los alumnos. *Formación Universitaria*, 2, 3-14. <http://dx.doi.org/10.4067/S0718-50062009000400002>
- Martínez-Gómez, M., Marín-García, J.A., & Girado-O'Meara, M. (2013). Validation of the JDS satisfaction scales applied to educational university environments. *Journal of Industrial Engineering and Management*, 7(1), 72-99. <http://dx.doi.org/10.3926/jiem.906>

- Mason, M.M. (2012). Motivation, Satisfaction, and Innate Psychological Needs. *International Journal of Doctoral Studies*, 7, 259-277.
- Millsap, E.R., & Yun-Tein, J. (2004). Assessing Factorial Invariance in Ordered-Categorical Measures. *Multivariate Behavioral Research*, 39(3), 479-515. [http://dx.doi.org/10.1207/S15327906MBR3903\\_4](http://dx.doi.org/10.1207/S15327906MBR3903_4)
- O'Neil, D.A., & Hopkins, M.M. (2002). The Teacher as Couch Approach: Pedagogical Choices for Management Educators. *Journal of Management Education*, 26, 402-414. <http://dx.doi.org/10.1177/105256290202600406>
- Paris, S.G., & Turner, J.C. (1994). Situated motivation. In Pintrich, P., Brown, D., & Einstein, C. (Eds.). *Student motivation, cognition, and learning: Essays in honor of Wilbert J. McKeachie*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Peinado, J.E., Chávez, A., Viciano, J., & Rivero, J.G. (2012). Invarianza factorial del cuestionario de autoeficacia EACA en universitarios de ciencias de la salud y ciencias sociales. *Formación Universitaria*, 5(4), 37-48. <http://dx.doi.org/10.4067/S0718-50062012000400005>
- Reinig, B.A., Horowitz, I., & Whittenburg, G.E. (2011). The Effect of Team-Based Learning on Student Attitudes and Satisfaction. *Decision Sciences Journal of Innovative Education*, 9, 27-47. <http://dx.doi.org/10.1111/j.1540-4609.2010.00289.x>
- Richardson, J.T.E. (2005). Instruments for obtaining student feedback: a review of the literature. *Assessment & Evaluation in Higher Education*, 30, 387-415. <http://dx.doi.org/10.1080/02602930500099193>
- Santos-Rego, M.Á., Godás-Otero, A., Lorenzo-Moledo, M., & Gómez-Fraguela, J.A. (2010). Eficacia y satisfacción laboral de dos profesores no universitarios: Revisión de un instrumento de medida. *Revista Española de Pedagogía*, 245, 151-168.
- Satorra, A., & Bentler, P.M. (1994). Corrections to test statistic and standard errors in covariance structure analysis. In von Eye, A., & Clogg, C.C. (Eds.). *Analysis of Latent Variables in Developmental Research*. Thousand Oaks, CA: Sage. 399-419.
- Satorra, A., & Bentler, P.M. (2001). A Scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507-514. <http://dx.doi.org/10.1007/BF02296192>
- Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, 18(4), 210-222. <http://dx.doi.org/10.1016/j.hrmr.2008.03.003>

- Steenkamp, J.B.E.M., & Baumgartner, H. (1998). Assessing measurement invariance in crossnational consumer research. *Journal of Consumer Research*, 25, 78-90. <http://dx.doi.org/10.1086/209528>
- Tessema, M.T., Ready, K., & Yu, W.C. (2012). Factors Affecting College students' Satisfaction with Major Curriculum: Evidence from Nine Years of Data. *International Journal of Humanities and Social Science*, 2, 34-44.
- Ullman, J.B., & Bentler, P.M. (2004). Structural Equation Modeling. In Hardy, M., & Bryman, A. (Eds.). *Handbook of Data Analysis*. SAGE. 431-458. <http://dx.doi.org/10.4135/9781848608184.n19>
- Ünal, C., & Inan, H.Z. (2010). Students' perceptions of a situated learning environment. *Procedia - Social and Behavioral Sciences*, 2, 2171-2175. <http://dx.doi.org/10.1016/j.sbspro.2010.03.301>
- Vandenberg, R.J., & Lance, C.E. (2000). A review and synthesis on the measurement invariance literature: Suggestions, practices and recommendations for organisational research. *Organizational Research Methods*, 3, 4-70. <http://dx.doi.org/10.1177/109442810031002>
- Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), 486-492. <http://dx.doi.org/10.1080/17405629.2012.686740>
- Van Saane, N., Sluiter, J.K., Verbeek, J.H.A.M., & Frings-Dresen, M.H.W. (2003). Reliability and validity of instruments measuring job satisfaction - a systematic review. *Occupational Medicine*, 53, 191-200. <http://dx.doi.org/10.1093/occmed/kqg038>
- Widaman, K.F., & Reise, S.P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance abuse domain. In Bryant, K.J., & Windle, M. (Eds.). *The science of prevention: Methodological advance from alcohol and substance abuse research*. Washington, DC: American Psychological Association. 281-324. <http://dx.doi.org/10.1037/10222-009>

