Association for Information Systems
AIS Electronic Library (AISeL)

SAIS 2009 Proceedings

Southern (SAIS)

3-1-2009

An Assessment of the Mathematics Information Processing Scale: A Potential Instrument for Extending Technology Education Research

R. Franklin Morris Jr.
frank.morris@citadel.edu

L. Allison Jones-Farmer

R. Kelly Rainer Jr.

An Assessment of the Mathematics Information Processing Scale: A Potential Instrument for Extending Technology Education Research

R. Franklin Morris, Jr.
The Citadel
frank.morris@citadel.edu

L. Allison Jones-Farmer
Auburn University
joneall@auburn.edu

R. Kelly Rainer, Jr.
Auburn University
rainerk@auburn.edu

ABSTRACT

Many argue that the United States is falling behind other countries in technology innovation. Some attribute this situation to ineffective education in the areas of math, science, and technology. Research using affective measures has provided evidence of links between student attitudes in math and technology education. With the aim of extending the research, this study examines the psychometric properties of the Mathematics Information Processing Scale[1] (MIPS). The MIPS uses both cognitive and affective measures to explore various dimensions of students' approaches to learning statistics and mathematics. The original study used exploratory factor analysis, while this study uses confirmatory factor analysis to revise the MIPS instrument. By combining both cognitive and affective measures in a single instrument, the MIPS offers the potential to contribute new research knowledge toward the goal of improving math and technology education.

Keywords

Education, math, technology, psychometric assessment, measurement instrument, structural equation modeling

INTRODUCTION

Over the past several decades, a number of mathematics instruments have been designed to measure the affective (i.e. emotional) component of an individual's beliefs, attitudes, values, and anxieties towards mathematics. However, according to Bessant (1997), while, "cognitive processes, in general, have been the subject of comprehensive theoretical discussion, the range of [cognitive] operational measures (or procedures) developed for use in mathematics education has not followed suit" (p. 841). In addition, while cognitive variables such as reasoning skills, spatial orientation, and mathematical ability have been a part of studies investigating mathematics achievement, individual differences in cognitive style, learning approach, and study strategies have, by comparison, received little attention in mathematics education research (Bessant, 1997). To address this research gap, the Mathematics Information Processing Scale (MIPS) was developed by Bessant (1997) with the goal of offering a single instrument capable of exploring the dual cognitive and affective dimensions of student oriented mathematics behavior.

The original 87 item MIPS instrument was tailored to mathematics-statistics education and includes items of both domain-specific and content/curriculum-independent cognitive processes involved in learning statistical material (Bessant, 1997). The combination of cognitive and affective measures in the MIPS is intended to examine features of students' learning strategies, beliefs about mathematics-statistics, metacognitive problem-solving skills, attentional focus, and cognitive worry. Bessant (1997) suggests that such information may assist the interpretation and remediation of student-centered learning barriers. Because research has found links between student attitudes in math and technology education (e.g., Igbaria and Parasuraman, 1989; Harrison and Rainer, 1992; Morris and Thrasher, 2008), lessons learned about improving student attitudes and performance in mathematics may also have application for improving technology education. Thus, the MIPS holds potential for extending technology education research, which may contribute knowledge to help with the challenge of boosting innovation in the United States.

BACKGROUND

In the original development study, Bessant (1997) reports that the MIPS instrument was constructed to measure approaches students employ (a) to learn mathematics or statistics content, (b) to develop strategies for problem-solving, and (c) to utilize cognitive-attentional resources during testing or evaluation. According to Bessant (1997), the 87 items in the MIPS were developed from theory and research concerning metacognitive problem solving (9 items), student learning approaches (56 items), and cognitive-attentional models of test anxiety (26 items). All item responses in the MIPS utilize a 5-point Likert scale ranging from 1 "not at all" to 5 "almost always." The original study used a total sample of 340 introductory-level statistics students. Through exploratory factor analysis (EFA) of the 87 items, five dimensions were identified and labeled by Bessant (1997) as Metacognitive Problem Solving (MIPS_1), Surface-Disintegrated Study (MIPS_2), Deep-Associative Study (MIPS_3), Performance Preoccupation (MIPS_4), and Strategic Study (MIPS_5). The MIPS_5 factor, Strategic Study, was described in the original study as having a complex, multidimensional nature, which prompted the following statement of caution from the author: "some revision or expansion of MIPS items [within the MIPS_5 Strategic Study factor] appears necessary to more precisely explore the meaning of this particular statistics learning dimension" (Bessant, 1997, p. 854).

In the original study, EFA was used with an oblique factor solution because of theoretical relationships between some elements in the MIPS. This is noteworthy because results of oblique solutions are often difficult to replicate in future studies due to the very sample specific nature of the oblique approach (Hair, Anderson, Tatham, and Black, 1998, p. 110). Cronbach's alpha coefficient was used in the original study to assess the internal consistency reliability estimates for the five MIPS factors with the following results: (MIPS_1) Metacognitive Problem Solving (α = .89); (MIPS_2) Surface-Disintegrated Study (α = .88); (MIPS_3) Deep-Associative Study (α = .869); (MIPS_4) Performance Preoccupation (α = .85); and (MIPS_5) Strategic Study (α = .72). Bessant (1997, p. 855) reported that, "the five MIPS factors correlate as expected with theoretically related measures" and that, "the intercorrelation pattern indicates both a multidimensional analysis of the MIPS and a moderate amount of differential construct validity among scores on its five constituent scales." Thus, the original MIPS development study concluded with the identification of five factors and the retention of all 87 original items.

CURRENT STUDY

The current study employed a sample of 608 undergraduate upperclassmen enrolled in one of two required courses in the College of Business at a large southeastern university. A review of the characteristics of the participants indicated that the sample was representative of junior-senior level business students at the university. Data collection involved the use of a multipart questionnaire to gather input from the university students. The questionnaire incorporated demographic information and items from a variety of scales, including the 87 item MIPS instrument as the focus of the current study. To offset any potential order effects among the various scales, the MIPS was the first instrument in half of the questionnaires and last in the other half. Distribution of the two different questionnaires was random among the students.

Analytic Strategy

The development and evaluation of measurement scales have traditionally relied upon one or more of the following analyses: coefficient alpha, item-total correlations, and EFA (Gerbing and Anderson, 1988). According to Churchill (1979), these traditional metrics are most appropriate for assessing the goodness of measurement inherent within a particular scale in the absence of theoretical underpinnings (i.e., exploratory research). In situations, such as the current analysis, where there exists a sufficient theoretical foundation for measurement and interrelationships between variables, more contemporary techniques such as confirmatory factor analysis (CFA) are a more appropriate means of statistical testing (Segars, 1997). Therefore, the analytic strategy for the current study incorporates an approach similar to frameworks that have been developed in the areas of marketing research (e.g. Anderson and Gerbing, 1988), education (e.g. Joreskog and Sorbom, 1989), psychology (e.g. Bentler, 1986), and management information systems (MIS) (e.g. Segars, 1997). Following the Segars (1997) framework, three primary events constitute the process of assessing unidimensionality and other aspects of measurement quality: (a) theoretical measurement modeling, (b) statistical measurement modeling, and (c) grounded respecification.

DATA ANALYSIS AND FINDINGS

Importantly for the present study, the first step in the analytic strategy, theoretical measurement modeling, should have been completed during the original Bessant (1997) development study. Since development details for the 87 items of the original MIPS instrument are not discussed in the original study, it must be assumed that appropriate steps were taken to ensure adequate content validity for the items. The second event in the analytic strategy, statistical measurement modeling, has the purpose of assessing the consistency (reliability) and accuracy (validity) of the items representing the constructs of interest. In Bessant's (1997) original development study, EFA was used with an oblique factor solution to empirically derive five correlated factors from the 87 items. In the current study, CFA with maximum likelihood estimation (MLE) is used to analyze a covariance matrix representing a measurement model of the original MIPS solution. All model parameters were estimated using AMOS 5.0.1 (Arbuckle, 2003).

Confirmatory Factor Analysis

Prior to the data analysis, appropriate steps were taken to evaluate potential violations of the assumptions required for CFA and no significant problems were identified. The data (N = 608) were randomly split into two groups: a sample for analysis (n = 304) and a hold-out sample for reassessment of any respecified model (n = 304). The statistical power for the analysis of the original MIPS measurement model was found to be adequate (i.e. > 0.80) for the sample size and degrees of freedom (n = 304, d.f .= 3644) (MacCallum, Browne, and Sugawara, 1996).

Recommended measures of model fit were considered in assessing the adequacy of the model (Bentler, 1986; Segars and Grover, 1993). Such fit indices generally include those that indicate absolute or relative proportions of the observed covariances explained by the model (e.g. GFI, CFI, IFI) and their counterparts that are corrected for the number of parameters, thus accounting for the complexity or parsimony of the model (e.g. RMSEA) (Byrne, 2001, pp. 79-88; Kline, 2005, pp. 137-145). Table 1 presents a summary of key goodness of fit indices for the original five-factor, 87-item measurement model. Although the evidence includes some mixed indications, it is clear that the overall group of fit indices overwhelmingly suggests an extremely poor fit of the original model to the data. Hair et al. (1998, pp. 653-661) provide guidelines for assessing model fit for these and other indices; while Byrne (2001, pp. 79-88) provides a discussion of these and other fit indices specific to the AMOS output.

χ^2	*df*	χ^2/df	*signif.*	GFI	CFI	IFI	RMSEA	Lo 90	Hi 90
6549.269	3644	1.797	0.000	0.660	0.616	0.621	0.051	0.049	0.053

Table 1. Goodness of Fit Indices: Original Measurement Model for the MIPS

Grounded Respecification

The third step in the analytic process, grounded respecification, is only required in the case of unsatisfactory measurement models. The purpose of the third step is to reconcile the statistical evidence with existing theoretical knowledge to create a new model, which can then be reassessed using statistical measurement modeling. It is important to note that any respecified model should, like the original, also be grounded in theory and/or should employ additional data collection or a hold-out sample for reassessment (Segars, 1997; Shook, Ketchen, Hult, and Kacmar, 2004). In the absence of such cautions, any improvements realized through model respecification may not represent an actual improvement in the empirical representation of underlying phenomena, but may simply be the result of statistical chance (Chin and Todd, 1995; Segars and Grover, 1993). Thus, for the current analysis, a hold-out sample was used in reassessing the respecified model.

Overall, the results of the confirmatory factor analysis of the 87 item measurement model indicated that the original model was not consistent with the data from the current study. Therefore, the measurement model was considered unsatisfactory and respecification was indicated. Using a grounded respecification approach, any indicator that did not achieve a reasonably large factor loading (i.e. absolute standardized value > 0.50) was considered to possess low reliability (i.e. item reliability < 25%), since three-quarters (i.e. 75%) of such an indicator's variance is unique and unexplained by the factor it is specified to measure (Kline, 2005, p. 73). Therefore, following Kline's (1998, p. 217) advice, "Indicators so orthogonal to the rest could be dropped from the model."

Because the elimination of any item from the measurement model can have a cascading effect on other items and thus the overall model, a conservative model respecification process was followed. The conservative process employed an iterative approach for cautiously eliminating items with low reliabilities and/or other problems (e.g. non-significant parameters). The procedure involved: (a) first eliminating items below a minimum level of standardized loading; (b) rerunning the revised measurement model; and (c) reassessing the fit, loadings, and misspecification indices. This process was repeated until there were no additional offending items for which deletion or respecification could be justified. The final respecified model retained four factors, twenty-four items, and two covariances.

A review of the fit indices for the respecified model revealed considerable improvement versus the original model. Table 2 presents a summary of the key goodness of fit indices for the respecified measurement model. In total, the fit indices for the respecified model indicated a significant and substantial improvement in overall fit versus the original model.

χ^2	*df*	χ^2/df	*signif.*	GFI	CFI	IFI	RMSEA	Lo 90	Hi 90
426.754	250	1.707	0.000	0.898	0.921	0.922	0.048	0.040	0.056

Table 2. Goodness of Fit Indices: Respecified Measurement Model (MIPS-R)

An examination of the critical ratios revealed that all of the estimates in the respecified model were significant. A search through the model parameters for other signs of offending estimates also revealed no problems with the respecified model. Taken as a whole, the results of the confirmatory factor analysis for the four factor, 24-item respecified measurement model (MIPS-R) suggested an acceptable fit to the data. Thus, further validation of the respecified model was indicated.

Table 3 presents a summary of the two measures of reliability that were calculated for each of the four factors in the MIPS-R respecified model: Cronbach's alpha coefficient, and composite factor reliability. The Cronbach alpha's and the composite measures both suggest that the four constructs of the respecified MIPS-R model possess adequate levels of reliability.

Factor Number →	MIPS-R_1	MIPS-R_2	MIPS-R_3	MIPS-R_4
Cronbach Alpha Reliability	0.820	0.867	0.705	0.740
Composite Reliability	0.812	0.852	0.682	0.667

Table 3. Reliability Measures: Respecified Measurement Model (MIPS-R)

CFA analysis of the respecified model revealed relatively large correlations between factors 1 and 3 ($r = 0.552$) and 2 and 4 ($r = 0.756$). However, as a rough guideline, the fact that these two correlations are less than 0.85 suggests discriminant validity for the factors (Kline, 2005, p. 73). An empirical basis for assessing discriminant validity is accomplished by creating a nested model under the original model and fixing the correlations between the factors equal to one. This nested model arrangement allows for a chi-square difference test on the values obtained for the constrained and unconstrained models. If the model in which the correlations are not constrained has a significantly lower chi-square value versus the model that is constrained, it would indicate that the factors are not perfectly correlated and that discriminant validity can be inferred (Segars, 1997). Such a test for the respecified model indicated that the factors were significantly distinct ($\Delta\ \chi2 = 201.272$, $p = 0.000$), and therefore inferred discriminant validity.

Validation of Respecified Model

To ensure that any improvement in measurement is not simply due to statistical chance, the respecified model was validated using a sample different from the sample used for the analysis (Chin, 1998; Kelloway, 1995). More specifically, the hold-out sample ($n = 304$) was used for the CFA validation of the respecified model. The results of the validation analysis using the hold-out sample are shown in Table 4 and indicate a slightly reduced fit to the data versus the analysis sample. However, the overall results using the hold-out sample are sufficiently similar to those obtained with the analysis sample to indicate support for the validation of the respecified model.

χ^2	*df*	χ^2/df	*signif.*	GFI	CFI	IFI	RMSEA	Lo 90	Hi 90
439.191	250	1.757	0.000	0.892	0.904	0.906	0.050	0.042	0.058

Table 4. Goodness of Fit Indices: Validation of Respecified Measurement Model (MIPS-R)

A chi-square difference test of the respecified model using both the analysis sample and the hold-out sample revealed no significant difference between the two samples ($\Delta\ \chi2 = 25.767$, $df = 20$, $p = 0.174$). This result indicates that the factor structure in the model can be considered invariant across the two samples and provides additional support for the validation of the respecified model.

DISCUSSION, IMPLICATIONS, AND RECOMMENDATIONS

The evidence derived from the current study suggests that the original five-factor model did not provide an adequate fit to the data, and that 63 of the original 87 items could be eliminated from the model. Such a radical reduction in the number of items may understandably raise questions about the model, the analysis, or both. However, a review of the original MIPS study provides additional insight and rational support for such a sweeping modification to the model. Two key findings from a review of the original MIPS study are shared here as follows.

First, is the finding that the five factors in the original MIPS model intrinsically possess multidimensional qualities as well as many items with low reliabilities. Since Bessant (1997), after reviewing the results of the EFA, did not eliminate any of these problematic items from the original 87 items in the original MIPS model, it is reasonable that the current study should identify such a large number of items for deletion. In addition, as forewarned by Bessant (1997), the MIPS_5 Strategic Study factor was found to possess the greatest number of problematic items, which led to the eventual elimination of the entire MIPS_5 Strategic Study factor from the model.

Second, is the realization that the complex factors in the MIPS make it more challenging for a researcher to define the constructs precisely, which can lead to problems with content validity. Therefore, while the respecified MIPS-R model offers a marginally acceptable instrument that provides cognitive-affective measures for mathematics-statistics evaluation, future studies should undertake a meticulous re-evaluation and fundamental reworking of at least the weaker items for each factor using a rigorous approach such as Segar's (1997) three-steps of theoretical measurement modeling. Such an exacting reassessment of content validity could provide a better measurement instrument with improved validity and reliability. In short, theoretically deriving and refining items during the initial stage of theoretical measurement modeling is considered critical for proper scale development and testing (Churchill, 1979; Gerbing and Anderson, 1988; Segars, 1997).

SUMMARY AND CONCLUSION

The original MIPS instrument combines cognitive and affective measures to examine facets of learning strategies for statistics and mathematics related content, metacognitive problem solving skills, and attentional deployment in evaluative contexts. Bessant (1997) is to be commended for his original development study, which takes an important first step in the construction of an instrument that combines cognitive and affective measures for mathematics-statistics evaluation. Beyond the initial development study, however, there are no other studies identified that have investigated the MIPS. Therefore, the present study extends the extant literature and contributes new information that future research may build upon in the continuing development of the MIPS instrument.

The results of the present study indicate that the respecified MIPS-R model provides a much more parsimonious instrument and an improved model fit for this sample. However, the evidence also suggests that respecification can only offer limited improvement and that more refinement is needed to further strengthen the reliability and validity of the MIPS-R. An important recommendation of the current study is that future research should undertake a rigorous and thorough reassessment of the content validity of the MIPS-R scales. Such a reassessment can lead to necessary improvement in the accuracy and consistency of the MIPS-R measures. Until then, it is prudent and practical to limit use of the MIPS-R to research purposes only.

The combination of cognitive and affective measures in the MIPS-R, "may assist the interpretation and remediation of student-centered learning barriers;" while, "group based administration of the MIPS-R may also provide important information concerning the nature of cognitive and affective factors operating in the instructional setting" (Bessant, 1997, p. 855). The instrument may also prove to be a valuable tool in further research exploring the links between mathematics and technology education (e.g., Morris and Thrasher, 2008). Such research may provide insights that may ultimately contribute to improving technology innovation in the United States. This study has advanced the MIPS a step forward in its development toward such worthy applications. Thus, further development of the MIPS-R is not only appropriate, but encouraged.

PLEASE NOTE: Due to space limitations, the authors have not included certain tables and figures that further support this research. Please contact the corresponding author to request any such information not included in this article.

REFERENCES

1. Anderson, J. C., and Gerbing, G. W. (1988) Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach, *Psychological Bulletin, 103*(3), 411-423.
2. Arbuckle, J. L. (2003) AMOS 5.0.1 (Version "Build 5151"), Chicago, IL: SmallWaters Corp.
3. Bentler, P. M. (1986) Structural Modeling and Psychometrika: An Historical Perspective on Growth and Achievements, *Psychometrika, 51*, 35-51.
4. Bessant, K. C. (1997) The Development and Validation of Scores on the Mathematics Information Processing Scale (MIPS), *Educational and Psychological Measurement, 57*, 841-857.
5. Byrne, B. M. (2001) *Structural Equation Modeling With AMOS* (1st ed.), New Jersey: Lawrence Eribaum Associates, Inc.
6. Chin, W. W. (1998) Issues and Opinion on Structural Equation Modeling, *MIS Quarterly, 22*(1), 7-16.
7. Chin, W. W., and Todd, P. A. (1995) On the Use, Usefulness, and Ease of Use of Structural Equation Modeling in MIS Research, *MIS Quarterly, 19*(2), 237-246.
8. Churchill, G. A. (1979) A Paradigm for Developing Better Measures of Marketing Constructs, *Journal of Marketing Research, 16*(February), 64-73.
9. Gerbing, D. W., and Anderson, J. C. (1988) An Updated Paradigm for Scale Development: Incorporatiing Unidemensionality and Its Assessment, *Journal of Marketing Research, 25*, 186-192.
10. Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1998) *Multivariate Data Analysis* (5th ed.), New Jersey: Prentice-Hall.
11. Igbaria, M., and Parasuraman, S. (1989) A path analytic study of individual characteristics, computer anxiety, and attitudes toward microcomputers, *Journal of Management,15*(3), 373-388.
12. Joreskog, K. G., and Sorbom, D. (1989) *LISREL 7 User's Reference Guide*, Chicago, IL: Scientific Software.
13. Kelloway, E. K. (1995) Structural Equation Modeling in Perspective Introduction, *Journal of Organizational Behavior, 16*, 215-224.
14. Kline, R. B. (1998) *Principles and Practice of Structural Equation Modeling*, New York: The Guilford Press.
15. Kline, R. B. (2005) *Principles and Practice of Structural Equation Modeling* (Second ed.), New York: The Guilford Press.
16. MacCallum, R. C., Browne, M. W., and Sugawara, H. M. (1996) Power Analysis and Determination of Sample Size for Covariance Structure Modeling, *Psychological Methods, 1*(2), 130-149.
17. Morris, R. F., and Thrasher, E. H. (2008) An Empirical Investigation of the Relationships among Math Confidence, Computer Confidence, and Computer Self-Efficacy: Implications for Technology Education *Proceedings of the Fourteenth Americas Conference on Information Systems*, August 14-17, Toronto, Ontario, Canada.
18. Segars, A. H. (1997) Assessing the Unidimensionality of Measurement: a Paradigm and Illustration Within the Context of Information Systems Research, *Omega, 25*(1), 107-121.
19. Segars, A. H., and Grover, V. (1993) Re-Examining Perceived Ease of Use and Usefulness: A Confirmatory Factor Analysis, *MIS Quarterly, 17*(4), 517-525.
20. Shook, C. L., Ketchen, D. J., Hult, G. T. M., and Kacmar, K. M. (2004) An Assessment of the Use of Structural Equation Modeling in Strategic Management Research, *Strategic Management Journal, 25*, 397-404.