

Fitting a Mixture Rasch Model to Visual Sequential Processing Memory Sub-dimension of ASIS: The Role of Covariates

Murat Doğan Şahin*

Received : 9 June 2020
Revised : 13 October 2020
Accepted : 17 December 2020
DOI : 10.26822/iejee.2021.190

*Correspondance Details: Murat Doğan Şahin.
Anadolu University, Faculty of Education,
Department of Educational Sciences, Eskisehir, Turkey.
E-mail: mdsahin@anadolu.edu.tr
ORCID: <http://orcid.org/0000-0002-2174-8443>

Abstract

Advanced Item Response Theory (IRT) practices serve well in understanding the nature of latent variables which have been subject to research in various disciplines. In the current study, 7-12 aged 2536 children's responses to 20-item Visual Sequential Processing Memory (VSPM) sub-test of Anadolu-Sak Intelligence Scale (ASIS) were analyzed with Mixture Rasch Model (MRM). In the first phase of the study, concomitant (covariate) variables were not used. In the second phase, age and gender were added to the model, and then the two models were compared in terms of fit indices, the number of latent classes and the distribution of item difficulties in the latent classes. The results of the study suggested that there were three latent classes in both models; however, the latter model had a better fit compared to the former model. In addition, the latent classes in both models had similar characteristics, and the distributions of item difficulties in the latent classes were also quite similar in both models while they had some differences in some aspects. The sizes of identical latent classes in both models varied between 15% and 30%. The results of the current study are expected to provide a deeper insight to researchers studying measurement theory and/or intelligence measurement.

Keywords:

Mixture Rasch Model, Concomitant/Covariate Variable, Latent Class, Visual Processing Memory

Introduction

New methods in measurement theory have a pivotal role in understanding the nature of latent variables which have been subject to research in various disciplines. In parallel with this view, new methods in measurement theory have been used in the measurement of intelligence which has a background of more than a century. Particularly advances related to Item Response Theory (IRT) continuously have offered critical advantages both theoretically and practically.

The sharpest advantage of IRT, when compared to Classical Test Theory (CTT), is the principle of parameter invariance. With IRT, ability prediction independent of items and item parameter prediction independent of groups are ensured (Embretson & Reise, 2000; Hambleton & Swaminathan, 1985).



Copyright ©
www.iejee.com
ISSN: 1307-9298

On the other hand, assuming the population from which item and ability parameters are predicted as a single homogenous population is a notable limitation for prediction through IRT. In fact, that a population consists of homogenous unknown sub-groups is a manifestation of this limitation (von Davier & Rost, 2017). This paves the way for new practices in IRT.

Another limitation for conventional IRT practices surfaces in bias studies (such as Differential Item Functioning-DIF). Although items providing advantage systematically to one of the groups of same ability level, or biased items in other words, can be identified with DIF studies, this practice is based on the assumption that the related groups are homogenous within themselves in terms of variables subject to measurement. Yet, it is known that individuals under the same manifest variable may consist of heterogeneous sub-groups (Samuelsen, 2005). Therefore, for whom items are biased cannot be identified and it cannot be understood for which reasons individuals respond differentially to items (Cohen & Bolt, 2005). This state comes along as another notable limitation for DIF practices based on manifest variables.

It can be argued that Mixture Rasch Model (Rost, 1990), which emerged through combining Rasch Model and latent class approach, offers a solution to these aforementioned limitations. Mixture Rasch Model (MRM) can be considered as a combination of latent class approach and IRT models (Frick et al., 2015). Accordingly, a continuous latent trait and latent class membership are predicted synchronously (Jiao et al., 2011). To put it another way, latent classes that are homogenous in itself but differ from other classes are identified, and group specific items and ability parameters are predicted in conjunction for these latent classes. Thus, it is postulated that invariance assumption is ensured for each latent class (Şen & Cohen, 2019). Compendiously, MRM combines classic Rasch models and latent class analysis and thereby it can exert item and ability parameters prediction for homogeneous sub-groups. This treatment provides identification of items showing DIF based on these latent classes, as well.

Another advantage of MRM is the use of concomitant variables. In the identification of latent classes, the effect of manifest covariate variables, also referred to as concomitant variables, on the formation of models can be tested, and which of these manifest variables have a notable contribution to the model can be revealed.

Within the scope of the current study, 20-item Visual Sequential Processing Memory (VSPM) sub-test of Anadolu-Sak Intelligence Scale (ASIS) was analyzed

with MRM and how the item parameters in the latent classes differ was examined. In the second phase of the study, the participants' genders and ages were added to the model as covariates and the contribution of these manifest variables to the model was tested.

Visual Sequential Processing Memory

VSPM is one of the three sub-tests measuring memory capacity in ASIS. These tests focus on processing memory and short-term memory. The theoretical framework of processing memory model in ASIS is based on Baddeley (2012) aiming to measure visual memory bandwidth. It is known that processing memory correlates at a high level with learning and academic achievement. A number of performances including basic reading, comprehension, mathematical calculation and reasoning are dependent upon processing memory capacity (Alloway, 2009; Dehn, 2014). Visual processing memory has an active role in the development of particularly basic mathematical abilities in early ages (Geary, 2011). Therefore, it can be asserted that studies on processing memory can provide significant implications regarding academic abilities of individuals in age groups for which the measurement is performed.

The items in the sub-test are mostly geometric and focus on sequencing of different shapes. Participants are provided with a sequence of shapes for a few seconds and then they are expected to pick it out among other sequences. There are fewer shapes in the beginning of the test; however, the number of shapes increases through the end. Using some shapes for more than once, various shape patterns are formed in some sequences (ASIS Manual, 2016).

Rasch & Mixture Rasch

Rasch Model (Rasch, 1960), a member of IRT family, makes use of only difficulty parameter while defining a relationship between individuals' ability levels and their likelihood of responding to a binary item. Accordingly, the logit of the difference between a person's ability and item difficulty provides responding likelihood of the item. These parameters are in interval scale. To determine the starting point in this scale, a constant reference point is identified in a way that difficulty of an item or total difficulty of all items is zero (Fischer, 1995). Given a person's ability subject to measurement is θ_i and item difficulty is β_j , response y_{ij} given by person i for item j is modeled as following (Rasch, 1960):

$$P(Y_{ij} = y_{ij} | \theta_i, \beta_j) = \frac{\exp\{y_{ij}(\theta_i - \beta_j)\}}{1 + \exp\{\theta_i - \beta_j\}} \quad (1)$$

Mixture model is a general approach used in order to model data which are thought to originate from different groups yet when membership to the group is not known. This modeling is as follows:

$$f(y_i) = \sum_{k=1}^K \pi_k f_k(y_i) \quad (2)$$

$f_k(\cdot)$ components in the formula may be densities or regression models. Mixture Rasch emerges with the combination of Formula 1 and Formula 2.

$$f(y|\pi, \psi, \theta) = \prod_{i=1}^n \sum_{k=1}^K \pi_k h(y_i|r_i, \beta_k) \psi_{r_i}, k. \quad (3)$$

In this model, also known as saturated model, a number of parameters that are not of concern in reality need to be predicted. Therefore, Rost and von Davier (1995) suggested a more parsimonious model based on only mean and variance when the number of items is more than four. In line with this, General Rasch Mixture Model turns out as:

$$f(y|\alpha, \theta, \psi) = \prod_{i=1}^n \sum_{k=1}^K \pi(k|x_i, \alpha) h(y_i|r_i, \beta_k) g(r_i|\psi_i) \quad (4)$$

The components in the formula are as follows:

$\pi(k|x_i, \alpha)$: Concomitant model for class membership,

$h(y_i|r_i, \beta_k)$: Class-specific likelihood of item difficulties,

$g(r_i|\psi_i)$: Class-specific score distribution.

Use of concomitant variables affecting prior class membership is also possible in MRM, and whether adding these variables provides a notable fit in the model or not can be tested. Concomitant variables model predicts the mixture and the influence of covariates simultaneously, which stands out as a notable advantage in contrast to other approaches to reveal the relationship between class membership and covariates ex post (Frick et al., 2012).

Research Purpose and Significance

In the first phase of the present research study, the distributions of item difficulties were examined according to latent classes that were formed with respect to VSPM levels of 7-12 aged individuals. It was also attempted to reveal the nature of latent classes based on the difference in this distribution. In the second phase, the effect of adding covariates on model fit and the distribution of the item difficulties in the latent classes in this second model (concomitant model) were examined, and this distribution was compared with the distribution of difficulties in the first model. Accordingly, MRM analysis, which bears the same purpose with latent class analyses yet, which is a much more robust and relatively new method, was performed in the study.

A number of studies are extant in the literature aiming to predict ability and item parameters within the scope of IRT models in different aspects of intelligence

tests (e.g., Beaujean & Osterlin, 2008; Ferreira et al., 2012; Schleicher-Dilks, 2015) or revealing bias through DIF practices (e.g., Abad et al., 2004; Colom et al., 2004; van der Sluis et al., 2008). On the other hand, it is also overt that Mixture IRT models are a better fit as opposed to conventional IRT models in cases when populations are not homogeneous (Muthén & Asparouhov, 2006). In this context, considering that a significant purpose of intelligence tests is to identify gifted individuals and categorize individuals in terms of related dimensions of intelligence test battery, it can be argued that the results to be obtained in the current study through MRM analysis would prove significant for the literature on measurement of intelligence.

VSPM is closely related to basic mathematics and language abilities of individuals in the 7-12 age groups. Herewith, the results of the study would also provide noteworthy findings for the researchers studying academic achievement of individuals in this age group. Additionally, this study is the first research study to examine the nature of latent classes that are formed through performing MRM analysis to an intelligence test and differentiation of item parameters among these classes.

Another purpose of the present study was to identify whether adding covariate variables to the model would cause a significant improvement in the model or not. To this end, ages and genders of the students who took the test were added to the model and thereby it was aimed to reveal the effect of these variables on both latent classes and obtained model parameters.

The current study sought to answer the following research problems. As a result of the MRM analysis;

- 1) How many latent classes are there in the first model? How are item difficulties distributed in terms of latent classes? In which context do the latent classes differ according to these distributions?
- 2) How do fit indices of the concomitant model which is formed through adding gender and age variables to the first model as covariates differ from the first model?
- 3) How are item difficulties distributed in concomitant model in terms of latent classes? In which aspects does this distribution differ from the first model or in which aspect are they similar?

Method

Participants

The data set used in this study was obtained from an implementation of Turkey norm study of ASIS, in which 4561 4-12 aged students from different regions of

Turkey selected through considering statistical region units identified by Turkish Statistical Institute took the intelligence test. Within the scope of the current study, data of 2536 7-12 aged students, of whom the youngest attended to elementary education, were used out of the data in Turkey norm study of ASIS.

Data Collection Tools

In this study, responses to 20 items belonging to VSPM dimension of ASIS were analyzed. ASIS (2016) is an intelligence test with seven sub-dimensions, developed to be implemented with 4-12 aged individuals. The number of total items in ASIS is 256. The responses given to 20 items belonging to VSPM dimension, used in this study, were obtained from norm study of ASIS.

Data Analysis

To answer the research questions, MRM analysis was performed in the first phase using responses of 2536 students to 20 items. In this step, the analyses were run without using any manifest variables. It is recommended to exclude participants who responded correctly or incorrectly to all of the items from the analysis because they do not contribute to conditional likelihood of item parameters (Rost, 1990). Therefore, eight students who did not have any correct answers were excluded from the study and the analyses were performed with the data of 2528 participants.

In Mixture IRT models, like latent class analysis, there is not a priori exact decision regarding the number of groups to emerge. For this reason, a vector is formed for potential latent class number. This vector is composed of whole numbers starting from 1 and continuing till a potential maximum value in order to identify the optimum number of latent classes with an exploratory approach. The optimum number of latent classes is identified for the model with the help of fit statistics obtained for each latent class number. Since latent class models formed in MRM are not nested in each other, information-based indices are used to evaluate models. Model-data fit increases as these values decrease. Of these indices, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Integrated Completed Likelihood Criterion (ICL) were used within the scope of this study. It is recommended in the literature that when there are contradictions among them, BIC values which are less biased should be used (Cubaynes et al., 2012). Accordingly, the optimum number of latent classes was identified through considering BIC values primarily. In the next step, the distributions of estimated item difficulties for each latent class were examined, and in which

contexts these latent classes differed was evaluated based on these distributions.

In the second phase of the study, the concomitant model was obtained through adding the participants' ages and genders to the model as covariate variables, and it was tested whether this new model had significant difference from the first model, which enabled the researcher to identify which of the manifest covariates could provide additional information. In the last phase, the similarities and differences of latent classes which emerged in the first and second models were put forth through examining item difficulty distributions in both models. The maximum number of latent classes was determined as four for both models.

The MRM analyses were run with psychomix (v1.1-8; Frick et al., 2012) package defined in R (R Core Team, 2020) with the method based on mean and variance of score distribution as suggested by Rost and von Davier (1995).

Results

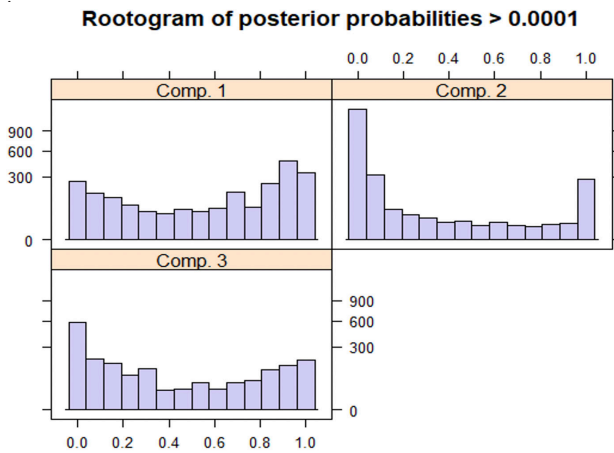
Within the scope of the study, first, BIC values were obtained in order to decide on the number of latent classes for the first model in which manifest variables were not used (see Table 1). Since the lowest BIC value was obtained for three-class solution, it was decided that the optimum latent class number for the data set was three.

Table 1
Fit Statistics for Different Number of Classes

Model	Number of Classes	Fit Indices		
		AIC	BIC	ICL
Model 1 (no covariate variables)	One class	39885.27	40007.81	40007.81
	Two classes	39105.26	39356.17	40171.90
	Three classes	38613.89	38992.03	39811.67
	Four classes	38693.62	39086.01	40131.71
Model 2 (Concomitant Model)	One class	39885.27	40007.81	40007.81
	Two classes	39004.16	39266.75	40085.41
	Three classes	34998.89	35395.68	36105.01
	Four classes	38442.66	38845.29	39517.65

With respect to VSPM, the participants were separated into three homogenous sub-groups. The histogram graph was also created to see posterior probabilities aiming to evaluate the three-component model (see Figure 1). In the U-shaped graph, the posterior distributions of the observations are seen as low or high, which reveals that the components (classes) are separated well enough at an acceptable level.

Figure 1
 Rootogram of posterior probabilities in the 3-component (class) MRM on VSPM data



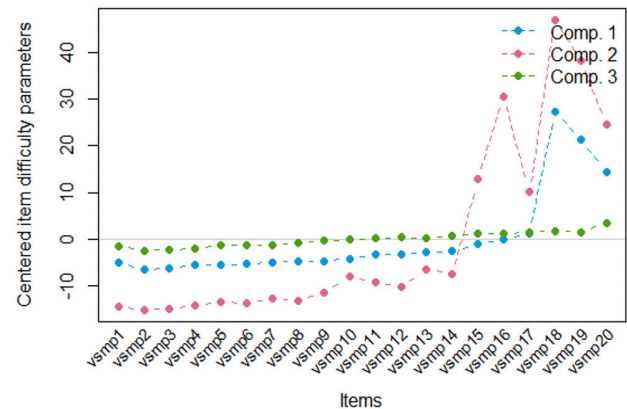
In the next step, the distributions of item difficulties for each latent class were obtained (see Figure 2). In the interpretation of the results, attention should be paid that the items are sequenced from simple to hard according to CTT item difficulties. In the examination of the obtained values, it is seen that item difficulty parameters of three classes are in parallel to a certain degree. Regarding the difficulties belonging to Class-2 (Component-2) shown in red, it can be argued that the first 14 items are perceived as easier for this class compared to other classes, and there is a relatively linear graph until this item. It catches attention that there are some bends between 10th and 14th items and there is a general increase in the difficulty values of the items. Starting with the 15th item, there are dramatic changes in difficulty values. Accordingly, it can be suggested that the last six items are very difficult for Class-2. The obtained difficulty values reveal that very few participants in this class could answer these items correctly.

With regard to difficulty parameters of Class-1 (Component-1), shown in blue, it is observed that the difficulty values of first part of the items for this class are between the values for the other two classes. In addition, it is seen that the values for Class-1 are in high parallelism with item difficulty parameters of Class-3. Such that, the difficulty levels of these two classes are very close to each other in the 15th and 16th items, and it is predicted as almost the same for the 17th item. However, starting with the 18th item, the case in the last three items is very similar to the case in Class-2, and there is a dramatic change in item difficulties. That the values obtained in Class-1 and Class-2 for the last three items are parallel to each other stands out. Finally, the difficulties obtained for Class-3 (component-3) were evaluated. The initial

items have a difficulty level slightly below 0, and item difficulty levels are in a slightly increasing fashion as the number of items increase. Accordingly, it can be suggested that the item difficulties of Class-3 form a line graph with a small slope. What is remarkable for Class-3 is that there are not dramatic changes in the graphs, and the highest value of the item difficulties is 3.5

Following a detailed examination of the graphs, latent classes were named. Considering that the difficulty values obtained for all three classes are different from each other but the item parameters follow a similar path until certain items, as also highlighted in the interpretation of graphs, it can be deduced that the classes were formed according to the participants' VSPM levels. That the classes were formed in line with VSPM levels is also supported by the fact that in Class-1 and Class-2, there are dramatic increases in item parameters as the item difficulty level increases, which is higher in Class-2, and the fact that item difficulties form a graph which could be labeled as linear in Class-3. Accordingly, since there are dramatic increases in item difficulties in the last six items, it is considered that Class-2 is the group which has the lowest VSPM level, and Class-1, in which a similar case emerges in the last three items, consists of medium-level individuals. Finally, it can be stated that Class-3, in which item difficulty indices vary at a very narrower range compared to the other groups and in which extreme values are not observed even in the items that can be described as very difficult (The highest difficulty value obtained is 1.9 except for the last item with the value of 3.5), is the group with the highest VSPM level.

Figure 2
 Item profiles for the 3-class MRM (Model-1) on VSPM data



In the second phase, the variables of gender and age were added to the model as covariate variables to answer the second and third research questions.

In the examination of fit indices (see Table 1), it was observed that, as the case in the first model, the lowest values were also obtained with three latent classes in this concomitant model which was formed through adding the manifest covariates. Therefore, it was deemed appropriate to compare three latent classes in both models. As mentioned beforehand, in mixture models, likelihood ratio test cannot be used in order to decide on the number of components. However, a comparison between the two models could be carried out with this test because the first model with three latent classes in which covariate variables were not used was nested in the second (concomitant) model which was also formed with three latent classes. The likelihood ratio test which yields a test statistics of 1809.93 ($p < .001$), reveals that the concomitant model have a much better fit than the first model. In other words, the covariate variables improved the model significantly.

Another property of psychomix (Frick et al., 2012) package is that it can identify which of the covariates that formed the concomitant model have more effect on the formation of latent classes. As a result of the analysis to this end, it is observed that both covariates have effect; however, the absolute effect of age is greater than gender.

In the final phase of the study, the distribution of item difficulties obtained for latent classes in the concomitant model was examined (see Figure 3), and in which aspect this distribution differed from the first model was revealed.

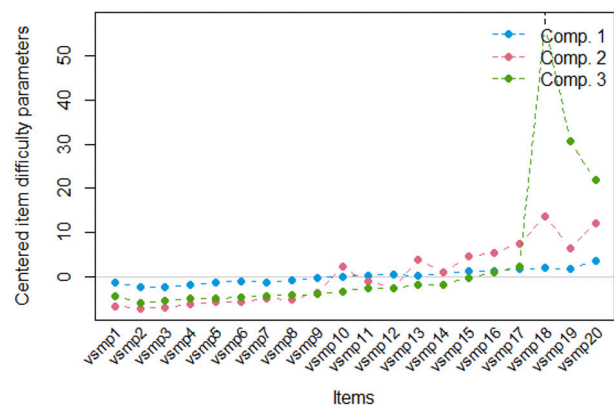
In this distribution, as in the first model, it is seen that the difficulty levels belonging to the initial items, in particular, are parallel to each other. However, differently from the first model, the item difficulty levels of the classes are much closer to each other. In particular, it is observed that item difficulty levels of Class-3, shown in green, and Class-1, shown in red, are very close to each other for the first ten items. It is also observed that the difficulty values of Class-2 have a rippled pattern starting with the 10th item and they get very high values starting with the 15th item. The horizontal pattern followed by the difficulty values of the initial items and the change in 10th-14th items lend their support to the fact that Class-2 is identical with the Class-2 in Model-1. On the other hand, it is also evident that the great change in the item difficulties in the first model starting with the 15th item is also observed in the concomitant model; however, this change is relatively less aggressive in the latter model. It is possible to say that Class-1 and Class-3 are identical with Class-3 and Class-1 in the first model, respectively. Accordingly, similar to the first model, the difficulty levels of these two classes progress in

parallel with each other, they get closer to each other and they get almost the same value in the 17th item. It is observed that there are dramatic increases in the item difficulty values for the last three items in Class-3, as is for the Class-1 which is its counterpart in the first model. Class-1 has a very similar distribution with Class-3 which is its identical in the first model.

The results of the current study indicate that the classes were formed in line with VSPM levels in a similar way with the first model. The number of the classes is the same and these classes are similar to their counterparts in the first model with respect to various characteristics. In particular, for Class-1, which consists of students with high VSPM level, very similar graphs were obtained in both models. The item difficulty levels of Class-2, involving students with the lowest VSPM level, started with higher values compared to the first model. It was also observed that the changes seen in the final items were not as dramatic as the changes in the first model. The difficulty distributions obtained in the class which included students with medium VSPM level were similar to the distributions in the first model; however, the amount of increase in the difficulty values of the last three items (differently from the first model) were much more higher than the values of Class-1.

To compare the first and second models, lastly, the numbers of individuals in the classes in the first and second models were compared. Accordingly, it was revealed that the class in the concomitant model which consisted of individuals with low VSPM levels expanded by 28% compared to the first model. The class in the concomitant model which consisted of individuals with medium VSPM levels expanded by 32% compared to the first model. Finally, the class consisting of individuals with high VSPM levels was examined and it was found out that it decreased by 15% compared to the first model.

Figure 3
Item profiles for the 3-class MRM (Model-2) on VSPM data



Discussion and Conclusion

Within the scope of the current study, VSPM sub-test of ASIS battery was analyzed with MRM. In the first phase of the study, no covariate variables were used. The number of latent classes was decided as three and it was evaluated that the latent classes which were obtained through their difficulty distributions were formed according to the individuals' VSPM levels as low, medium and high. In the second phase, the concomitant model was formed through adding gender and age variables as covariates, optimum number of classes was found as three as well, and it was observed that the covariate variables improved the model significantly compared to the first model. Of these variables, it was revealed that age had a higher contribution to the model and gender had a relatively lower contribution. In the second model, similar to the case in the first model, the latent classes were formed according to the individuals' VSPM levels as low, medium and high. The comparison of the item difficulty distributions of identical classes in both models suggested that they were mostly similar yet some items differed to some degree. Accordingly, after adding covariate variables to the model, changes were observed in the distribution of item difficulties of some items. Besides, it was concluded that the membership number of classes that had similar characteristics in both models changed in the range of 15% and 30%.

The results obtained in this study were first examined with respect to the use of covariate variables in MRM analysis. It was highlighted in some former studies that adding manifest covariates contributed to identification of latent classes in Mixture IRT models, revealing the differences among these classes (e.g., Choi et al., 2015; Li et al., 2016), and prediction of parameters (Dai, 2013). The findings in the present study overlap with some of these studies. On the other hand, it was observed that the manifest covariate variables used in this study had an effect on class membership to some degree but they did not cause a change in the number of classes. This finding is in agreement with those obtained by Karadavut et al. (2019).

The finding of the current study that age and gender had a high degree of contribution to the model, gender with a greater degree, is in line with those of previous studies in the literature. A number of studies unearth that there is increase in individuals' visual processing memory and tasks related to it until the ages of 11-12 (Brockmole & Logie, 2013). Heyes et al. (2016) observed that visual processing memory improved precision in middle childhood. There are other studies in the literature lending support to this finding (Cowan et

al., 2010; Cowan et al., 2011). Voyer et al. (2017), who put forth the relationship between visual processing memory and gender with a comprehensive meta-analysis study, reported that visual processing memory differed significantly in terms of gender yet this was at a very low level.

The studies in the literature broadly lend their support to the effect of covariate variables on the model identified in the current study. Yet these studies predominantly depend on CTT based measurements. In the studies aimed at measuring intelligence, as in the current study, the use of MRM analysis may contribute more to understanding the nature of the characteristic subject to measurement.

The present study focused on visual processing memory which is a critical aspect of intelligence. On the other hand, this would be a fruitful area for further work. Future studies may exert similar procedures for the other aspects of intelligence. It may be interesting to identify the changes in the model when different manifest variables other than gender and age or latent covariates are added, and compare these changes with findings of the current study.

The factors affecting the development of children's basic reading and basic mathematics abilities have always been a major area of interest for researchers. The current study dwelled on a factor that is known to be closely related to basic reading and mathematical abilities. Further research should be undertaken to explore different aspects of intelligence that may be closely related to children's academic achievement.

References

- Abad, F. J., Colom, R., Rebollo, I., & Escorial, S. (2004). Sex differential item functioning in the Raven's Advanced Progressive Matrices: Evidence for bias. *Personality and Individual Differences*, 36(6), 1459–1470. [https://doi.org/10.1016/S0191-8869\(03\)00241-1](https://doi.org/10.1016/S0191-8869(03)00241-1)
- Alloway, T. P. (2009). Working memory, but not IQ, predicts subsequent learning in children with learning difficulties. *European Journal of Psychological Assessment*, 25(2), 92–98
- ASIS Manual. (2016). Anadolu Sak Zeka Ölçeği (ASIS) Uygulayıcı Kitabı. Anadolu Üniversitesi ÜYEP Merkezi, Eskişehir.
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1-29.

- Beaujean, A. A., & Osterlind, S. J. (2008). Using Item Response Theory to assess the Flynn Effect in the National Longitudinal Study of Youth 79 Children and Young Adults data. *Intelligence*, 36(5), 455–463. <https://doi.org/10.1016/j.intell.2007.10.004>
- Brockmole, J. R. & Logie, R. H. (2013). Age-related change in visual working memory: A study of 55,753 participants aged 8–75, *Frontiers in Psychology*, 4.
- Colom, R., Escorial, S. & Rebollo, I. (2004). Sex differences on the Progressive Matrices are influenced by sex differences on spatial ability. *Personality and Individual Differences*, 37(6), 1289–1293. <https://doi.org/10.1016/j.paid.2003.12.014>
- Cohen, A. S. & Bolt, D. M. (2005). A mixture model analysis of differential item functioning. *Journal of Educational Measurement*, 42, 133–148.
- Choi, Y. J., Alexeev, N. & Cohen, A. S. (2015). Differential item functioning analysis using a mixture 3-parameter logistic model with a covariate on the TIMSS 2007 mathematics test. *International Journal of Testing*, 15, 239–253.
- Cowan, N., Morey, C. C., AuBuchon, A. M., Zwillling, C.E. & Gilchrist, A. L. (2010). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Developmental Science*, 13, 120–133.
- Cowan, N., AuBuchon, A. M., Gilchrist, A. L., Ricker, T. J. & Saults, J. S. (2011). Age differences in visual working memory capacity: Not based on encoding limitations. *Developmental Science*, 14(5), 1066–1074. doi:10.1111/j.1467-7687.2011.01060
- Cubaynes, S., Lavergne, C., Marboutin, E. & Gimenez, O. (2012). Assessing individual heterogeneity using model selection criteria: How many mixture components in capture–recapture models? *Methods in Ecology and Evolution*, 3, 564–573.
- Dai, Y. (2013). A mixture Rasch model with a covariate: A simulation study via Bayesian Markov chain Monte Carlo estimation. *Applied Psychological Measurement*, 37, 375–396.
- Dehn, M. J. (2014). *Essentials of processing assessment*. 2. Edition. Wiley.
- Embretson, S. E. & Reise, S. P. (2000). *Item Response Theory for psychologists*. Erlbaum.
- Ferreira, A. I., Almeida, L. S. & Prieto, G. (2012). Construction of a memory battery for computerized administration, using Item-Response Theory. *Psychological Reports*, 111(2), 585–609.
- Fischer, G. H. (1995). Derivations of the Rasch Model. In Fischer & Molenaar (Ed.), *Rasch Models foundations, recent developments, and applications*, (pp. 15–38).
- Frick, H., Strobl C., Leisch F. & Zeileis A (2012). Flexible Rasch Mixture Models with package psychomix. *Journal of Statistical Software*, 48(7), 1–25. <http://www.jstatsoft.org/v48/i07/>.
- Frick, H., Strobl, C. & Zeileis, A. (2015). Rasch mixture models for DIF detection: A comparison of old and new score specifications. *Educational and Psychological Measurement*, 75(2), 208–234.
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: a 5-year longitudinal study. *Developmental Psychology*, 47(6), 1539–1552. <https://doi.org/10.1037/a0025510>
- Hambleton, R. K. & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Kluwer Academic Publishers.
- Heyes, S. B., Zokaei, N. & Husain, M. (2016). Longitudinal development of visual working memory precision in childhood and early adolescence, *Cognitive Development*, 39, 36–44.
- Jiao, H., Lissitz, R. W., Macready, G., Wang, S. & Liang, S. (2011). Exploring levels of performance using the mixture Rasch model for standard setting. *Psychological Test and Assessment Modeling*, 53(4), 499–522.
- Karadavut, T., Cohen, A. S., & Kim, S. H. (2019). Mixture Rasch model with main and interaction effects of covariates on latent class membership. *International Journal of Assessment Tools in Education*, 6(3), 362–377.
- Li, T., Jiao, H. & Macready, G. B. (2016). Different approaches to covariate inclusion in the mixture Rasch model. *Educational and Psychological Measurement*, 76(5), 848–872.
- Muthén, B. & Asparouhov, T. (2006). Item response mixture modeling: Application to tobacco dependence criteria. *Addictive Behaviors*, 31, 1050–1066. doi:10.1016/j.addbeh.2006.03.026

- R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. The University of Chicago Press.
- Rost, J. (1990). Rasch Models in latent classes: An integration of two approaches to item analysis. *Applied Psychological Measurement*, 14(3), 271-282.
- Rost, J. & von Davier M. (1995). *Mixture distribution Rasch Models*. In Fischer & Molenaar (Ed.), *Rasch Models foundations, recent developments, and applications*, (pp. 257-268).
- Sak, U., Bal-Sezerel, B., Ayas, B., Tokmak, F., Özdemir, N.N., Demirel-Gürbüz, Ş. & Öpengin, E. (2016). *Anadolu Sak Intelligence Scale: ASIS Practitioner's Book*.
- Samuelsen, K. M. (2005). *Examining differential item functioning from a latent class perspective*. [Unpublished doctoral dissertation]. University of Maryland, College Park.
- Schleicher-Dilks, S. (2015). *Exploring the item difficulty and other psychometric properties of the core perceptual, verbal and working memory subtests of the WAIS-IV using Item Response Theory*. [Unpublished Doctoral Dissertation]. Florida: NSU.
- Şen, S. & Cohen, A. S. (2019) Applications of mixture IRT models: A literature review. *Measurement: Interdisciplinary Research and Perspectives*. 17(4), 177-191, doi: 10.1080/15366367.2019.1583506
- van der Sluis, S., Posthuma, D., Dolan, C. V., de Geus, E. J. C., Colom, R. & Boomsma, D. I. (2006). Sex differences on the Dutch WAIS-III. *Intelligence*, 34, 273–289.
- von Davier M. & Rost, J. (2017). Logistic mixture-distribution response models. In W. J. van der Linden (Ed.), *Handbook of item response theory, volume one: Models* (p. 393-406). Chapman and Hall.
- Voyer, D., Voyer, S.D. & Saint-Aubin, J. (2017). Sex differences in visual-spatial working memory: A meta-analysis. *Psychon Bull Rev* 24, 307–334. <https://doi.org/10.3758/s13423-016-1085-7>