Validation of Cognitive Dimension of Managerial Aptitude Test: Rasch Model Analysis

Rita Markus Idulfilastri[1*] Meylisa Permata Sari[1] Cynthia Sutanto[1]

[1] *Faculty of Psychology, Universitas Tarumanagara, Jakarta 11440, Indonesia*
[*] *Corresponding author. Email: ritamarkus@fpsi.untar.ac.id*

ABSTRACT
Differentiated Model of Giftedness and Talent (DMGT) explains that every person has innate potentials that can be transformed into talent [1]. The initial construction of Managerial Aptitude Test developed to measure the three dimensions: cognitive, metacognitive, and creativity, using *Classical Test Theory* (CTT) approach and shows satisfactory result. The current study further examined the psychometric properties of Managerial Aptitude Test using one-parameter *Item Response Theory* (IRT), Rasch Model. Cognitive dimension was chosen from the three dimensions of Managerial Aptitude Test based on the dynamics of thinking process in decision making among the research sample. Rasch model analysis performed to measure item difficulty level in term of participants ability. Total of 413 employees from three different types of company (Government, State-owned, Private) participate in the study. The result of Rasch analysis supports the unidimensional of 25 items in Cognitive Dimensions. Person-Item map shows same logit value = 0. In term of item difficulty level, 3 items have high difficulty level where none of the participants able to answer the items. On the other hand, 3 items also found to be very easy, therefore every participant (include participant with low potential) able to answer those items correctly. Item validity indicates the need to eliminates item 1, 14, and 15. *Differential Item Functioning* result shows several items contains bias.
Keywords: *cognitive ability, managerial aptitude test, Rasch model*

1. INTRODUCTION

The concept of Differentiated Model of Giftedness and Talent (DMGT) explains that every person has innate potentials, a gift that can be transformed into talent [1], and this transformation manifested as an ability. The transformation process is supported by learning activities and practical experiences. Gagne [1] further explains that natural potential appears in the form of intelligence, creativity, social and physical ability; while talent can be seen from academic abilities, artistic abilities, business skills, skills to related with enjoyment, social skills, technology skills, and movement abilities. Previous literature review study [2] suggests that managerial potential as natural potential consists of three dimensions of latent construct: cognitive, metacognitive, and creativity.

Managerial potential reflected through competency in problem solving and decision making, ability to achieve work target, and also ability to convey ideas. By identifying managerial potential, Idulfilastri [2] suggest that career achievement can be predicted. From the result of their study using employee as research sample, [3] argues that coworkers' assessment significantly predicts employee performance by controlling personality traits. Thus, Law et al. [3] findings suggests that managerial potential is not the sole predictors of competencies, but still, the role of managerial potential needs to be considered. Therefore, an instrument is needed to measure managerial potential.

Few studies conducted to develop instrument to measure the potential, for example study by Idulfilastri [4]. Using classical test theory (CTT) approach, Factor Analysis methods, Idulfilastri [4] found 17 valid items to measure cognitive dimension of managerial potential. However, several studies have pointed out weaknesses of CTT approach [5], for examples: (1) subject's ability and item difficulty are not considered, (2) the score analysis is a raw score containing error, (3) level of difficulty and item weighting are highly dependent on the characteristics of the group. Therefore, in the current study we test the cognitive dimension of Managerial Aptitude Test using Rasch model.

Rasch analysis allowing probability approach toward measured attributes, use true score therefore free from measurement error, overcomes item weight differences, and mapping subject's ability and item difficulty with a same unit.

1.1. Related Work

Cognitive dimension in Managerial Aptitude Test conceptualized using the critical thinking approach by Watson and Glaser [6]. Critical thinking defined as the dynamics of individual thinking by generating

assumptions, reflect on the thoughts, performs analytical synthesis, evaluates and identify logical relationships. Several characteristics attributed to critical thinker, such as: questioned certain assumptions in order to make correct inferences, reflect on one's own thoughts in order to understand information in a correct manner, perform a synthesis analysis of some of the results of thoughts and evaluate the existing information, and perform reasoning of all information to generates a conclusion that can be communicated appropriately [6].

Several validation studies of Watson Glaser Critical Thinking Appraisal (WGCTA) [7,8,9,10] indicates that WGCTA has robust psychometric properties. In earlier study, Behren [7] found WGCTA score positively correlates with student's GPA (r=0.51-0.56), result that also shared in a study by Williams [8] where WGCTA score shows positive correlation with mid-term exam (r=0.42) and final exam scores (r=0.57). The findings suggest evidence for criterion validity of WGCTA. Another result of validation study using 273 students from a private university in Lebanon, show evidence for unidimensional of WGCTA. WGCTA score also shows positive and significant relationship with GPA (*r=0.41*), TOEFL (*r=0.47*), SAT-Verbal (r=*0.43*), SAT-Math (*r=0.33*), and Academic Proficiency (CAAP) Critical Thinking Test (r=0.64), indicates the evidence of criterion validity of the instrument [9]. In similar vein, findings in Gadzella, Stacks, Stephens, and Masten [10] study shows WGCTA as a robust instrument to measure critical thinking for students pursuing career as a teacher, where WGCTA score positively correlates with critical thinking score (*r=0.31*). WGCTA however, measure in general population rather for certain position that requires higher level of critical thinking.

In recent study involving 322 employees at managerial level position from various types of company backgrounds (financial, construction, IT consultant etc.), Idulfilastri [2] developed Managerial Aptitude Test. The measure later validated using CFA second order with estimate true score. Psychometric result shows model fit for every item in cognitive dimension (χ^2*=117.75*, *df*=95, *P-value*=0.06, *RMSEA*=0.027), which indicates an empirical fit to the theoretical concept, with 19 valid items [4]. The validation, however, relies on CTT. In current trend, psychometric assessment using IRT approach which generally generates more accurate ability measurement than using CTT.

1.2. Our Contribution

While Managerial Aptitude Test shows satisfactory psychometric property in the initial validation study [4], the limitation of using CTT highlighted. Thus the current paper extend the information of psychometric properties of Managerial Aptitude Test: Cognitive Dimension based on one parameter of Item Response Theory (IRT): item difficulty using Rasch analysis.

1.3. Paper Structure

The paper is organised as follows. Section 2 served as methods section, which include the characteristics of research participants, measures used, and the description of how the data will be analysed. Section 3 consists of statistical analysis result of Rasch model using Winstep and discussion of the result. In Section 4, we concludes the paper and suggestion for future research.

2. METHODS

2.1. Participants

Participant in the study selected using purposive sampling technique, based on the consideration of those in the managerial level or able to occupy managerial positions at higher level, and are determined by the company to participate in assessment centre. 413 manager level participant reported here gathered from three different companies: Government (*n* = 103, 24.94%), State-Owned (*n* = 112, 27.12%), and Private (*n* = 198, 47.94%). Most of the participant in this study were male (*n* = 343, 83.05%), with less than a quarter of female participants (*n* = 70, 16.95%).

2.2. Managerial Aptitude Test: Cognitive

Based on psychological test construction, Managerial Aptitude Test classified as an aptitude test, a power test with a choice of correct and wrong answers [11][12][13]. Cognitive ability in this test operationalised as thinking skills, an ability that enable individual to perform critical thinking in problem solving and decision making. Cognitive dimension consists of five indicators: (a) inference, (b) assumption, (c) deduction, (d) interpretation, and (e) evaluation of arguments. Participants are assigned to answer 25 items by choosing right or wrong based on story or statement given. Correct answer coded as 1, and 0 for the wrong answer.

Two examples given in original language (Indonesian language).

Item 1. "Story: Pada hari Kemerdekaan RI, banyak pejuang yang menghadiri upacara pengibaran bendera merah putih. Bila dilihat dari peserta yang hadir ternyata banyak yang berusia lanjut. Kemungkinan di antara mereka ada yang mengikuti langsung peristiwa kemerdekaan pada tahun 1945. Butir pertanyaan: Upacara pengibaran bendera merah putih dilaksanakan di Istana Negara RI". Answer Right or Wrong.

Item 2. "Question: Pada hari Senin banyak pekerja berangkat lebih pagi dibandingkan hari lainnya. Butir pertanyaan: Pada hari Senin pagi, lalu lintas lebih macet dibandingkan hari-hari lainnya". Answer Right or Wrong.

2.3. Analysis Technique

Rasch model analysis performed to produces fit statistics which to indicates the extent of the data obtained ideally illustrates that the participant with higher cognitive ability provide patterns of responses to items according to their level difficulty [5]. Information weighted fit (infit) and outlier sensitive fit (outfit) of the mean square and standardized value used as parameters. According to Sumintono and Widhiarso [14], infit measures the sensitivity of response pattern to the items targeted on the person; while outfit measures the sensitivity of response pattern to items with certain difficulty level on the person.

Rasch analysis is run using Winsteps [15]. There are two information considers to determine whether or not the data fit to the Rasch model: (a) the mean square is 1.0, with z-standardized values = 0.0. In item – person level analysis, there are three parameters requires to determines item's fit-or-infit: (1) *Point Measure Correlation* (x) : $0.32 < x < 0.8$; (2) *Outfit Mean Square* (y): $0.5 < y < 1.5$; (3) *Outfit Z standard* (z): $-2.0 < z < +2.0$. Infit item is one that considered as "too easy" (high negative logit value) or "too difficult" (high positive logit value) based on the response pattern given by participants; or the value of three criteria generates from the analysis shows the item does not fit the requirements, which indicates the item does not measure the desired characteristics [14].

The analysis stages in this paper are described as follows: (1) Unidimensional. The first requirement that must be met for using Rasch model is that the dimension has unidimensional properties, meaning that it only measures one attribute [15]. Principal Component Analysis (PCA) of the residuals in Rasch Analysis model are used to determine the extend to wish the variability of the instrument measuring one attribute that should be measured; (2) Reliability: Response and Item. This statistical analysis used to determine the relationship between person ability and items difficulty using fit statistics. The parameters used are the infit and outfit mean square error (MNSQ) and standardized values based on the response pattern to the items [14]. Values of Infit MNSQ shows the sensitivity of response pattern to the target items on participant, while Outfit MNSQ measures the sensitivity of the response pattern to items with certain difficulty level for the participant [5]; (3) Validity. The validity of Managerial Aptitude Instrument examined to measure the magnitude of the attributes measured by testing the validity of the construct and validity of the content, while the validity of respondents was to investigate inconsistent response patterns of participants. Person-Item Map allowed us to examine the distribution of participant's response pattern (person validity) and level of approval of the items (item validity). For person validity, if the map shows item logit and person logit has a same logit average (0.0), the participant's average ability is equivalent to the standard difficulty level of the item [14]. Item validity determine by the average value of the items. Average item of 0.0. logit represents probability of error and 50% probability of correctness between level of ability of participant with item difficulty level [16]; and (4) Differential Item Functioning (DIF) result shows whether certain items are function in different ways for participant with certain characteristics and inspect irrelevant factors [17], where certain item prefer one particular type of participant's characteristic. DIF analysis result with probability value below 5% (0.05) indicates significant bias on the item.

3. RESULT AND DISCUSSION

3.1. Unidimensional

Table 1 shows the result of PCA of Rasch residuals in Winsteps. The raw variance explained by the data is 26.7%, not much different from the expected value (26.8%). As suggested by Reckase [18], the total raw variance explained by this measure is 26.7% (above 20%), supported the unidimensional of cognitive dimension. Another evidence of unidimensional also showed by the unexplained variance (10%) which indicates the level of item independence in cognitive dimension is categorized as good. Thus, based on the variance, expected value, and unexplained variance, cognitive dimension fulfill the unidimensional assumption.

Table 1 Cognitive dimension: Standardized residual variance (in Eigenvalue unit)

```
TABLE 23.0 COGNITIVE                                ZOU997WS.TXT  Jun 14 16:51 2020
INPUT: 413 Person  25 Item  REPORTED: 413 Person  25 Item  2 CATS  WINSTEPS 3.73
-------------------------------------------------------------------------------

     Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
                                            -- Empirical --     Modelled
Total raw variance in observations     =         34.1 100.0%          100.0%
  Raw variance explained by measures   =          9.1  26.7%           26.8%
    Raw variance explained by persons  =          1.2   3.5%            3.5%
    Raw Variance explained by items    =          7.9  23.3%           23.3%
  Raw unexplained variance (total)     =         25.0  73.3% 100.0%    73.2%
    Unexplned variance in 1st contrast =          2.4   6.9%   9.4%
    Unexplned variance in 2nd contrast =          1.5   4.4%   6.0%
    Unexplned variance in 3rd contrast =          1.4   4.2%   5.8%
    Unexplned variance in 4th contrast =          1.4   4.1%   5.6%
    Unexplned variance in 5th contrast =          1.4   4.1%   5.6%
-------------------------------------------------------------------------------
```

3.2. Reliability: Person and Items

Rasch analysis were performed using Winsteps to process 413 data to 25 items in the cognitive dimension, the result as the following: *chi-square value* = 11252.88 with degree of freedom (*df*) =9888, with significant test probability (*p=0.000)* indicates a significant and satisfactory of overall measurement. Table 2 and Table 3 provide statistical summaries of Rasch analysis performed. In Rasch analysis, measures reliability obtained from infit and outfit result, as shown from Table 2 and Table 3. Mean person measure obtained =-0.03 logit, indicates participants' average score in cognitive dimension. Score lower than 0.00, suggest that most participants responded with a wrong answer. Reliability using Cronbach's measures the interaction between person and item in general shows poor person reliability (0.00), while the item reliability shows a good result (0.99), indicates poor response consistency however quality of test items is exceptional.

The participant's answer pattern can be measured by the MNSQ infit value and the MNSQ outfit with the expectation = 1.0. In Table 2, the person has an infit value of MNSQ is 1.00 and outfit MNSQ is 1.00, this shows that the overall response pattern of the response on the instrument is good. Whereas for the ZSTD infit value and the ZSTD outfit the expectation is = 0.0, this shows that overall the participant's answer pattern is in accordance with the model.

In Table 3 for items that have an infit MNSQ and outfit MNSQ = 1.00, shows that overall the items on the instrument are good. Whereas for the ZSTD infit value and the ZSTD outfit the expectation is = 0.0, this shows that overall the items of the Cognitive dimension have compatibility with the model. When the analysis is continued based on the identification of the respondent group and the grain group, the value of separation is used. By using the strata separation equation, it turns out that from the person group with separation = 0.00, the value is 0.33, meaning that there is only one group of participants.

Table 2 Cognitive dimension: Person-fit statistical summary

```
TABLE 3.1 COGNITIVE OU997WS.TXT  Jun 14 16:51 2020
INPUT: 413 Person  25 Item  REPORTED: 413 Person  25 Item  2 CATS  WINSTEPS 3.73
-------------------------------------------------------------------------------

     SUMMARY OF 413 MEASURED Person
-------------------------------------------------------------------------------
          TOTAL                          MODEL         INFIT        OUTFIT
          SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD
-------------------------------------------------------------------------------
 MEAN      12.3      25.0        -.03     .47      1.00     .0   1.00     .0
 S.D.       2.1         .0         .46     .01       .23    1.1    .33    1.1
 MAX.      19.0      25.0        1.50     .53      1.88    3.6   2.50    3.9
 MIN.       6.0      25.0       -1.50     .46       .47   -3.1    .41   -2.6
-------------------------------------------------------------------------------
 REAL RMSE    .49 TRUE SD     .00  SEPARATION   .00  Person RELIABILITY  .00
MODEL RMSE    .47 TRUE SD     .00  SEPARATION   .00  Person RELIABILITY  .00
 S.E. OF Person MEAN = .02
-------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .00
```

Table 3 Cognitive dimension: Item-fit statistical summary

```
-------------------------------------------------------------------------------
          TOTAL                          MODEL         INFIT        OUTFIT
          SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD
-------------------------------------------------------------------------------
 MEAN     203.4     413.0         .00     .12      1.00     .0   1.00     .0
 S.D.     100.5        .0        1.23     .02       .03     .7    .05     .9
 MAX.     371.0     413.0        2.06     .16      1.05    1.7   1.09    2.0
 MIN.      49.0     413.0       -2.29     .10       .94   -1.4    .90   -1.9
-------------------------------------------------------------------------------
 REAL RMSE    .12 TRUE SD    1.22  SEPARATION 10.12  Item   RELIABILITY  .99
MODEL RMSE    .12 TRUE SD    1.22  SEPARATION 10.17  Item   RELIABILITY  .99
 S.E. OF Item MEAN = .25
-------------------------------------------------------------------------------
```

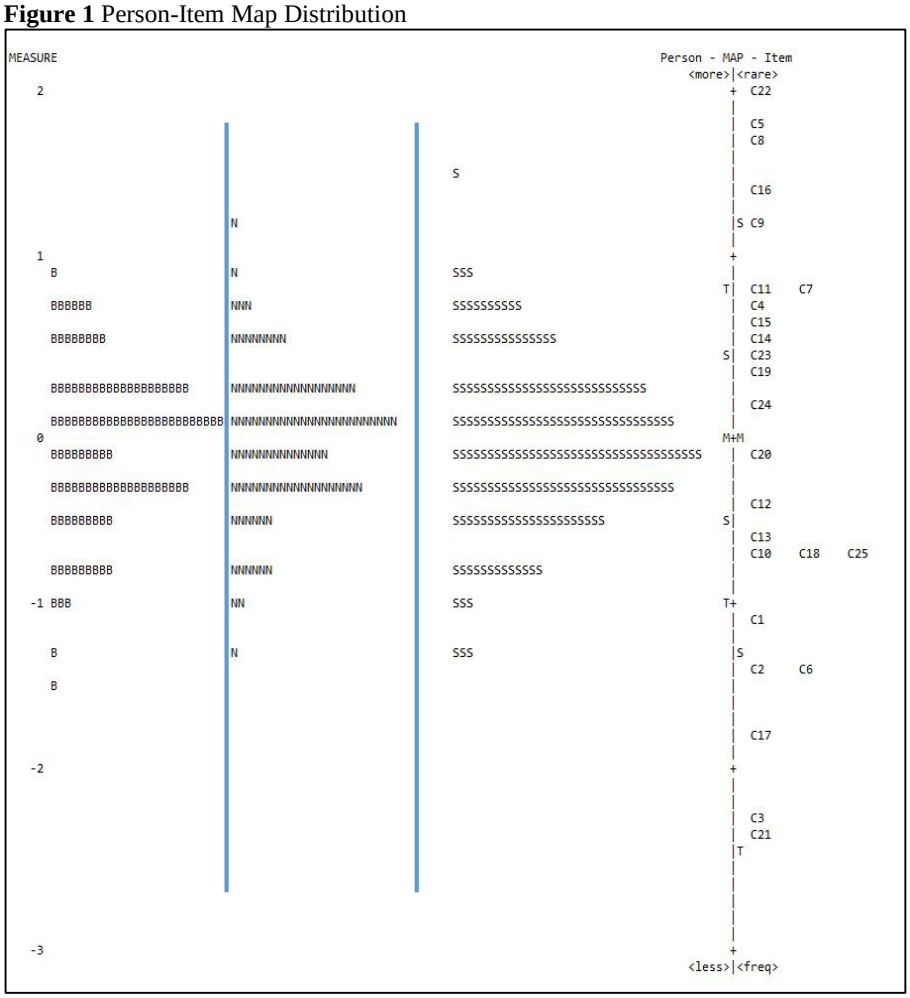
3.3. Validity

3.3.1. Person Validity

Figure 1 shows the distribution of response patterns (left side) and level of approval of the items (right side). The right side shows the grouping of participants based on the type of company, with N represents participants from Government, B represents state-owned employee, and S for participants from private company. Response pattern of participants based on workplace shows an interesting tendency, where participants from the three companies exhibit various response pattern from those who easily answer correctly (high positive logit value/top of the map), to respondents who found it difficult to answer correctly (high negative logit value/ the most bottom part of the map. Figure 1 shows that participant and item logit placed at the same value, 0.0, which means that the average cognitive of the participant is equivalent to the standard difficulty level of the items.

Result also shows that there are three items with high difficulty level, even higher than every participant's ability to answer correctly: C8, C5, and C22. In another word, no participants managed to answer the question correctly. On the other side, there are also three items (C17, C3, and C21) that a very low level of difficulty, where every participant in the study able to answer the question correctly, participants with low ability included.

Figure 1 Person-Item Map Distribution



rata-rata logit aitem 0,0 logit

rata-rata logit person 0,0 logit

3.3.2. Item Validity

Table 2 shows the item average value is 0.0 logit, indicate a satisfactory instrument validity, meaning that cognitive dimension of Managerial Aptitude Test able to measure participants' cognitive ability properly. On the left side of Figure 1, all items in the measurement (from the smallest to highest logit value) located within participant logit value, suggests that every item measures participants' cognitive ability.

Table 4 shows the item misfit order, where every item has positive value of point measure correlation, with standard error +0.12 logit. Therefore, item with infit MNSQ higher than 1.03 indicates misfit (C15, C1, C14). However, if using this criterion: 2<ZSTD<+2; 0,5<MNSQ<1,5; 0,4<Pt Mean Corr<0,85, all item is fit, meaning they have different response patterns therefore all items can be maintained. The arrangement of cognitive dimension is as follows: C15, C1, C14, C8, C17, C25, C3, C7, C19, C20, C13, C21, C18, C10, C9, C23, C22, C16, C12, C5, C2, C24, C4, C6, C11. However, for detailed analysis, item C15, C1, and C14 can be eliminated if necessary.

Table 4 Cognitive dimension: Item Fit

```
TABLE 10.1 COGNITIVE                                ZOU997WS.TXT  Jun 14 16:51 2020
INPUT: 413 Person  25 Item  REPORTED: 413 Person  25 Item  2 CATS  WINSTEPS 3.73
-------------------------------------------------------------------------------
Person: REAL SEP.: .00  REL.: .00 ... Item: REAL SEP.: 10.12  REL.: .99

         Item STATISTICS:  MISFIT ORDER

-------------------------------------------------------------------------------------------
ENTRY   TOTAL  TOTAL           MODEL|   INFIT   |  OUTFIT  |PT-MEASURE |EXACT MATCH|
NUMBER  SCORE  COUNT  MEASURE  S.E. |MNSQ  ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%  EXP%| Item
------------------------------------+----------+----------+-----------+-----------+------
    15    138    413     .70     .11|1.05   1.4|1.09   2.0|A .07   .21| 66.1  67.1| C15
     1    298    413   -1.02     .11|1.05   1.0|1.09   1.5|B .07   .20| 72.6  72.3| C1
    14    152    413     .54     .10|1.05   1.7|1.06   1.6|C .10   .21| 60.5  64.2| C14
     8     62    413    1.78     .14|1.01   .2|1.06   .6|D .11   .16| 85.0  85.0| C8
    17    348    413   -1.77     .14|1.00   .0|1.03   .3|E .16   .17| 84.3  84.2| C17
    25    271    413    -.70     .11|1.03   .8|1.02   .6|F .15   .21| 64.2  66.4| C25
     3    369    413   -2.23     .16|1.01   .1|1.03   .2|G .11   .14| 89.3  89.3| C3
     7    124    413     .86     .11|1.02   .5|1.02   .4|H .16   .20| 70.2  70.2| C7
    19    171    413     .34     .10|1.01   .4|1.02   .6|I .19   .22| 60.8  61.0| C19
    20    209    413    -.05     .10|1.01   .5|1.01   .5|J .20   .22| 60.5  59.0| C20
    13    261    413    -.59     .10|1.01   .3|1.01   .2|K .19   .21| 65.1  64.6| C13
    21    371    413   -2.29     .16| .99   .0|1.01   .1|L .15   .14| 89.8  89.8| C21
    18    269    413    -.68     .11|1.00   .1| .99  -.2|M .21   .21| 65.1  66.1| C18
    10    273    413    -.72     .11|1.00   .0|1.00   .1|l .21   .21| 68.0  66.8| C10
     9     91    413    1.30     .12| .99  -.1| .98  -.2|k .20   .18| 78.2  78.0| C9
    23    159    413     .47     .10| .98  -.5| .99  -.3|j .24   .21| 63.7  63.0| C23
    22     49    413    2.06     .15| .99  -.1| .95  -.4|i .18   .14| 88.1  88.1| C22
    16     78    413    1.50     .13| .99  -.1| .95  -.5|h .21   .17| 81.4  81.1| C16
    12    245    413    -.42     .10| .98  -.8| .98  -.7|g .26   .22| 63.9  62.0| C12
     5     60    413    1.82     .14| .98  -.2| .94  -.5|f .21   .16| 85.5  85.5| C5
     2    322    413   -1.35     .12| .98  -.3| .95  -.6|e .25   .19| 78.2  78.0| C2
    24    187    413     .17     .10| .97 -1.3| .96 -1.4|d .28   .22| 61.7  59.5| C24
     4    133    413     .76     .11| .97  -.8| .96  -.7|c .27   .20| 70.2  68.1| C4
     6    320    413   -1.32     .12| .96  -.7| .92 -1.0|b .30   .19| 77.2  77.5| C6
    11    124    413     .86     .11| .94 -1.4| .90 -1.9|a .35   .20| 69.7  70.2| C11
------------------------------------+----------+----------+-----------+-----------+------
 MEAN   203.4  413.0     .00     .12|1.00   .0|1.00   .0|           | 72.8  72.7|
 S.D.   100.5    .0    1.23     .02| .03   .7| .05   .9|           |  9.8   9.9|
-------------------------------------------------------------------------------------------
```

3.4. Differential Item Functioning

DIF analysis was performed to examine the test equivalence across different types of company. As suggested by previous studies [19] [20], DIF higher than 0.5 logits indicate that the item will be performed differently by participant from different types of company. Table 5 displays unbiased and biased items among three different types of company. The result affirms that C1, C3, C5, C8, C9, C13, C18, C19, C22 and C23 measuring cognitive dimension irrespective of different types of company. However, C2, C4, C6, C7, C10, C11, C12, C14, C15, C16, C17, C20, C21, C24, and C25 showed high DIF which leads to bias toward certain characteristics of the participant. It shows that the assessment of cognitive abilities influenced by various things which in the current analysis cannot be ascertained. A possible explanation is due to different types of companies in the study, which might contribute to variances in work experiences, educations, gender, and the values that govern employees.

Table 5 *DIF* types of company: Government, State Owned, Private

```
TABLE 30.4 COGNITIVE                              ZOU350WS.TXT  Jun 30 18:25 2020
INPUT: 413 Person  25 Item  REPORTED: 413 Person  25 Item  2 CATS  WINSTEPS 3.73
-------------------------------------------------------------------------------

DIF class specification is: DIF=$S5W1
```

Person CLASSES	SUMMARY DIF CHI-SQUARE	D.F.	PROB.	BETWEEN-CLASS MEAN-SQUARE	t=ZSTD	Item Number	Name
3	.7947	2	.6704	.1259	-1.1632	1	C1
3	9.2257	2	.0097	1.4825	.7541	2	C2
3	1.0842	2	.5789	.1419	-1.1021	3	C3
3	2.1587	2	.3362	.3973	-.4613	4	C4
3	.6192	2	.7329	.1056	-1.2485	5	C5
3	2.8953	2	.2319	.3808	-.4922	6	C6
3	5.7386	2	.0555	.8939	.2232	7	C7
3	.6660	2	.7157	.1161	-1.2029	8	C8
3	.3339	2	.8471	.0631	-1.4722	9	C9
3	2.9516	2	.2254	.4509	-.3662	10	C10
3	4.6636	2	.0953	.8846	.2132	11	C11
3	3.2445	2	.1945	.5270	-.2435	12	C12
3	.1514	2	.9289	.0273	-1.7634	13	C13
3	1.7223	2	.4191	.3070	-.6429	14	C14
3	5.4640	2	.0638	.7405	.0474	15	C15
3	2.1085	2	.3448	.2799	-.7042	16	C16
3	2.9152	2	.2296	.3983	-.4594	17	C17
3	.6465	2	.7228	.1142	-1.2111	18	C18
3	.0996	2	.9533	.0188	-1.8694	19	C19
3	4.9770	2	.0814	.7061	.0048	20	C20
3	2.6358	2	.2643	.4303	-.4017	21	C21
3	.9674	2	.6142	.1610	-1.0345	22	C22
3	.2051	2	.9042	.0309	-1.7255	23	C23
3	1.6530	2	.4340	.2771	-.7109	24	C24
3	2.1507	2	.3376	.3518	-.5489	25	C25

4. CONCLUSION

The result of Rasch analysis support a unidimensional cognitive dimension of managerial potential test, it also shows a satisfactory items reliability. The current result suggests that Cognitive Dimension of Managerial Potential Test valid at item level, however, improvement needed in term of person levels, which shows that response given are inconsistent with participant ability. The result implies respondent's ability to answer the items has yet measured in the same unit, and several items unable to identify respondent's ability. In terms of DIF, the result also suggests that several items contain bias based on the type of companies where the participant work. Considering satisfactory item quality, but weak person fit in the current study, and also several biased items, further research needs to examine the difficulty level of several items and consider whether or not certain items should be revise, replace or deleted.

ACKNOWLEDGMENT

This work was funded by Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Tarumanagara, Jakarta, Indonesia.

REFERENCES

[1] F. Gagne. Building Gifts into Talent: Overview of the DMGT. University Quebec, Montreal (Canada), 2008.

[2] R. M. Idulfilastri. Tes Potensi Manajerial Berbasis Ranah Kognitif untuk Memprediksi Kinerja Calon Karyawan, Doctoral. Dissertation, Universitas Indonesia, 2012.

[3] K. S. Law, C. S. Wong, L. J. Song. The construct and criterion validity of emotional intelligence and its potential utility for management studies. Journal of Applied Psychology. 89(3) (2004) 483-496. DOI: 10.1037/0021-9010.89.3.483

[4] R. M. Idulfilastri. Pengujian konstruk Tes Potensi Manajerial berdasarkan validitas butir dengan metode Factor Analysis. Jurnal Muara Ilmu Sosial, Humaniora, dan Seni. 2(1) (2018) 189-197. DOI: 10.24912/ jmishumsen.v2i1.1597

[5] B. Sumintono, W. Widhiarso. Aplikasi Pemodelan Rasch pada Assessment Pendidikan. Trim Komunikata, Cimahi, Jakarta, 2015.

[6] G. Watson, E. M. Glaser, Watson-Glaser: Critical Thinking Appraisal – Short Form Manual. Pearson Education, New York (NY), 2008.

[7] P. J. Behrens. The Watson-Glaser Critical Thinking Appraisal and Academic Performance. Journal of Nursing Education. 35 (1996) 34-36. https://pubmed. ncbi.nlm. nih.gov/8926515

[8] R. L. Williams. Critical thinking as a predictor and outcome measure in a large undergraduate educational

psychology course. Research Reports, University of Tennessee, 2003 (Also appears in ERIC ED478075)

[9] K. E. Hassan, G. Madhum. Validating the Watson Glaser Critical Thinking Appraisal. Higher Education. 54 (2007) 361-383. DOI: 10.1007/s10734-006-9002-z

[10] B. M. Gadzella, M. Baloglu, R. Stephens. Prediction of GPA with educational psychology grades and critical thinking scores. Education. 122(3) (2002) 618-623. https://www.researchgate.net/profile/Mustafa_Baloglu/publication/286926564_Prediction_of_GPA_with_educational_psychology_course_grades_and_critical_thinking_skills/links/567023f808ae2b1f87acd3cc.pdf

[11] R. M. Kaplan, D. P. Saccuzzo. Psychological Testing: Principles, Applications, and Issues. Fifth Edition. Wardsworth Thomson Learning, Los Angeles (CA), 2001.

[12] R. J. Cohen, M. E. Swerdlik. Psychological Testing and Assessment: An Introduction to Test and Measurement. Sixth Edition. McGraw-Hill, New York (NY), 2005.

[13] J. Mankar., D. Chavan. Differential Aptitude Testing of Youth. International Journal of Scientific and Research Publications. 3(7) (2013) 1-6. http://www.ijsrp.org/research-paper-0713/ijsrp-p1937.pdf

[14] B. Sumintono, W. Widhiarso. Aplikasi model Rasch untuk Penelitian Ilmu-ilmu Sosial. Trim Komunikata, Cimahi (Indonesia), 2013.

[15] J. M. Linacre. A User's Guide to WINSTEPS MINISTEPS, Rasch-Model Computer Programs. Winsteps.com, Chicago (IL), 2016.

[16] T. G. Bond, M. C. Fox. Applying the Rasch Model fundamental measurement in the human sciences. Third Edition. Routledge, New York (NY). 2015.

[17] B. D. Wright, M. H. Stone. Best Test Design. Mesa Press, Chicago (IL). 1979.

[18] M. D. Reckase. Unifactor latent trait models applied to multifactor tests: Results and implications. American Educational Research Association. 4(3) (1979) 207-230. http://www.jstor.org/stable/1164671?seq=1&cid=pdf-reference#references_tab_contents

[19] B. D. Wright, N. Panchapakesan. A procedure for sample-free item analysis. Educational and Psychological Measurement. 29 (1969) 23-48.

[20] W. C. Wang, G. Yao, Y. J. Tsai, J. D. Wang, C. L. Hsieh. Validating, improving reliability, and estimating correlation of the four subscales in the WHOQOL-BREF using multidimensional Rasch analysis. Quality of Life Research. 15 (2006) 607e620.