# PROFILING OF EDUCATION POLICY MAKERS' STATISTICAL LITERACY: A DIFFERENTIAL ITEM FUNCTIONING ANALYSIS OF DEMOGRAPHIC BACKGROUNDS

**Azlin Abd Jalal**[1]
**Harris Shah Abd Hamid**[2]
***Hutkemri Zulnaidi**[1]
[1] Faculty of Education, Universiti Malaya
[2] Faculty of Management, Education & Humanities, University College MAIWP International (UCMI)
**hutkemri@um.edu.my*

**Abstract:** In this data-driven society, the need to be statistically literate is increasing in importance as reading and comprehending statistics is integral to making informed decisions. Statistical literacy is even more important for education policy makers who are responsible for policy outcomes which greatly impacts the national education sector. This study thus aims to measure the statistical literacy of education policy makers and assess which demographic characteristics showed possibility of bias in assessing statistical literacy using the Rasch model approach. Participants self-reported their statistical literacy with 20 items tailor made to the work of education policy makers. Data were drawn from a survey elicited using a cross-sectional survey on education personnel working at different levels in Ministry of Education. The respondents consist of 328 selected personnel in a proportionate cluster random sampling. The main assumptions of the Rasch model were satisfied. Analysis showed that there are no bias items in this study, however three bias items were flagged, where one subgroup in the demographic characteristics showed better ability to answer the items compared to another. Profiling of the policy makers provided insights on area of strength and weaknesses within level of statistical literacy based on their demographic characteristics for statistical training needs and planned professional development programs, recruitment and placement for education policy makers.

***Keywords:*** Education Policy Makers, Statistical Literacy, Profiling, Differential Item Functioning

## INTRODUCTION

Statistical literacy is beneficial for almost all teachers, leaders, and policy makers in educational environments as they strive to better understand students' learning and try to improve every aspect of learning outcomes and school improvement (Henderson & Corry, 2020). According to Wallman (1993, p.1), the heart of statistical literacy is understanding statistical information which will result with better decision making, thus, defining statistical literacy as "the ability to understand and critically evaluate statistical results that permeate our daily lives and also the ability to appreciate the contribution that statistical thinking can make in public and private, professional and personal decisions." Gal (2002) further explained that statistical literacy is not just being able to read and evaluate data and graphs but also to involve the elements of critical evaluation and disposition such as attitude and belief especially in social contexts to understand media reports. The highlights on statistical literacy have elevated in the past few decades, as statistical literacy construct evolves in tandem with social change. This is more pertinent in the current pandemic situation where education policy makers were needed to make drastic decisions on education such as investigations on unprecedented school closures, calculation of learning loss and skill loss as well as its' impact on the nations' economy which needed to be handled instantly with statistical literacy (Budgett & Rose, 2017; Hanuschek & Woessmann, 2020; Watson & Callingham, 2020).

The education system is the most essential foundation in developing a stronger society as betterment of education provides better economic and social benefits to the nation (UNESCO, 2018). In the education sector, as in any other workplace settings, accountability, quality control, and forward planning could be enlightened by exploration of statistical data (Chick & Pierce, 2012). The education policy makers are held accountable for all policy decisions made regarding school improvement, resource allocation, curriculum planning and also intervention programs (Pierce et al., 2014; Reeves & Chiang, 2018; Sharma, 2017; United Nations Economic Comission for Europe, 2012). Thus, they are expected to better understand statistics published in their workplace such as statements, media reports, research and statistical reports including the national and international assessments. In the Malaysia Education Blueprint 2013-2025, it is mentioned that there is a limited use of data for informed decision-making although one of the most capital-intensive investment by the ministry is the data collection and management systems (Ministry of Education, 2013). Moreover, the key findings from the Survey

of Adult Skills showed that 20% to 40% of adults in most countries have low or very low numeracy skills (OECD, 2013) and even statistics majors were found to have less than 40% table and graph literacy (Tiro et al., 2018).

Although various studies were conducted on statistical literacy, almost all were intended for students, thus insufficient structured materials and empirical research were conducted for adult users especially education sector's stakeholders to better understand official statistics published (Gal & Ograjensek, 2017). Hence, this study aims to investigate the profile of statistical literacy among education policy makers based on demographic factors such as age, gender, tenure of service, and service institution in different levels in Ministry of Education, Malaysia. In addition, this study aims to explore the subsequent differential item functioning to assess scale variations to investigate the possibility of bias in items assessing statistical literacy.

## LITERATURE REVIEW

The underlying theory for the study is the Theory of Cognitive Constructivism (Piaget, 1972) on how humans translate the interaction between their environmental experiences and their biological maturation ideas. Life experience includes knowledge, attitudes, and beliefs, in relation to their ability to understand and interpret statistical information. The model of statistical literacy as proposed by Gal (2002) stated that knowledge elements and other facilitating processes need to be present for a statistical literate person to understand, critically evaluate, interpret, and communicate with statistical statements. The model encompassed two components which are the knowledge elements and another enabling processes which is known as the dispositional elements. The knowledge elements consist of literacy skills, mathematical knowledge, statistical knowledge, context knowledge and critical questions while the dispositional elements encompass of beliefs and attitude, and critical stance.

A substantial amount of literature has been published on statistical literacy. These studies touched on many areas in statistical literacy in various educational settings including studies on school students (Callingham & Watson, 2017; Maki & Horita, 2018; Sharma, 2017; Yolcu, 2014); postgraduate and adult learners (Kaplan & Thorpe, 2010; Schield, 2016); a survey on pre-service teachers and data analysts (Chick et al., 2014; Reeves & Chiang, 2018) and a survey on public officers in Philippines (Reston, 2005, 2010). In Malaysia, previous research done on statistical literacy tend to examine university students' (Akanmu & Jamaludin, 2015; Noor Lide et al., 2010; Krishnan, 2014). The motivation to assess adult statistical literacy among working professionals are provided by the challenges in the workplace related to data management and the scarcity of information and comprehensive empirical investigation concerning the education decision makers' statistical literacy aroused interest as it would be beneficial for both the organization and at the end, for the benefit of all students in Malaysia.

## METHODOLOGY

### *Research Design*

The study employs an instrument developmental method involving a quantitative research approach aiming to develop, evaluate and then, assess the statistical literacy of the respondents. A cross-sectional quantitative survey method was employed in this study using the self-developed instrument. This method will enable the researcher to collect data in a disperse population of education policy makers on their statistical literacy.

### *Participants*

Proportionate cluster random sampling technique was used to select 328 from the population of 2251 of education policy makers working from different levels of federal, state and districts working in Kuala Lumpur, Putrajaya, and Selangor. These respondents took the SL-EdP assessment, and the collected data were analysed using WINSTEPS version 4.8.1, a Rasch model measurement software. The education policy makers in this study consisted of 37.2% male and 62.8% female with 3.7% of them in the lowest age group of 21 and 30 years old, 46% in the 31 to 40 years old group, 38.7% others belong to the 41-50 age group while the rest 11.6 % are between 51 years and 60 years old. On the highest academic qualifications of the policy makers', almost half are PhD holders, 40.5% were Masters' degree holders, and only 12.8 % were bachelor's degree holders. With regards to grade, only one respondent (0.3%) is from the highest grade, JUSA, 1.2% from DG54, 10.4% from DG52, 32.9% from DG48, and the majority 51.8% were from grade DG44 while only 3.45 were from DG41. In terms of work experience, only 3.7% of the respondents have experience less than 10 years, most of them (46%) had tenure in service for 10-19 years, 38.7% of them have been working between 20 to 29 years, while another 11.6 % were already in service for 30 to 39 years. The education policy makers surveyed in this study involved

mostly of officers in various divisions of federal level MOE with 190 (57.9%); followed by 15.2% in the JPN and 26.8% from the PPD.

***Instrumentation***

A newly developed instrument named as the Statistical Literacy for Education Policy Makers Instrument (SL-EdP) was used to collect data in this study. This instrument was custom made to assess statistical literacy in their workplace where the items developed were loaded with issues and statistics regarding education such as dropouts, students' health, co-curriculum, and financial aid for poor students. Throughout the development of the instrument, evidence of reliability and validity were gathered. An instrument blueprint was first developed where a preliminary version of the item pool consisting of 30 items were reviewed by seven experts in the education field and statistics education. Additional rounds of revisions were then made in cognitive debriefing and a pilot test. Finally, the final version of SL-EdP with 20 multiple-choice items with 3 answer options was employed in the field study. The demographic part comprised of five basic demographic items including age, gender, service grade, working experience, and service institution while the statistical literacy assessment were 20 multiple-choice items with three answer options to measure the education policy makers' statistical literacy. The statistics topics assessed include data production (5 items), measures of distribution (2 items), descriptive statistics (4 items), graph (5 items), chart (2 items) and probability (2 items).

***Rasch Measurement Model***

In this study, statistical literacy was measured using SL-EdP instrument and Rasch measurement model was employed for data analysis as it was a mathematical model developed to allow for construction of instruments in measuring human latent traits in such a way based on a probabilistic relation between item difficulty with person ability (Wright & Masters, 1982). Furthermore, Rasch analysis allows for an approach which is unified to solve several issues of the measurement, including requirement for the validity of the transformation from sum raw score (ordinal) into interval scale, establishing the instrument's construct validity and unidimensionality, testing the item difficulties and the invariance of the items and whether there are bias in an item among subgroups of the respondents in differential item functioning (DIF) (Planinic et al., 2019). More precisely, this may indicate that the construct being measured has unidimensionality on a hierarchy based on the theoretical idealisation where the response pattern could be compared with the ones that do not coincide with the idea.

**RESULTS**

***Test of Validity and Reliability of the Instrument***

Prior to further analysis, validity and reliability tests were done to the developed the SL-EdP instrument. The scale was evaluated through Rasch analysis where construct validity was established by the item fit, point measure correlation, test of unidimensionality and the observation of the Wright map. Fit statistics indicated that all 20 items showed good fit. Item correlations also showed that the items were working in the same direction in measuring statistical literacy. In addition, test of unidimensionality also concluded that unidimensionality was evident as the factor in the residuals explained the most variance, thus supporting the existence of only one Rasch dimension while minimising possibility of any potential secondary dimension. Furthermore, the Wright map also showed a fairly good spread of person and items.

For the test of reliability, as illustrated in Table 1, the item reliability of .98 suggested that this instrument, SL-EdP could be confidently replicated to other samples and item separation index was 7.96, informing that the sample size in this study was large enough to establish the item difficulty hierarchy, thus affirming the construct validity. In addition, the person reliability estimate was 0.63 with person separation 2.08, indicating two to three distinguishable strata of education policy makers. Although the person reliability was not as good as the item reliability, the value is more than acceptable (Bond & Fox, 2015).

*Table 1*
*Test of Reliability and Validity of SL-EdP*

| | Test | | Analysis | Result |
|---|---|---|---|---|
| 1 | Construct Validity | i | Item fit | No item misfit |
| | | ii | Point Measure Correlations (PTMEA CORR) | No item less than 0.2 |
| | | iii | Unidimensionality Test | |
| | | a) | Variance explained (expected) | 30.6 (30.8) |
| | | b) | Unexplained variance in the first contrast | 7.4% |

| | | | | |
|---|---|---|---|---|
| | | c) | Eigenvalue | 2.1 |
| 2 | Reliability Test | i | Item reliability | 0.98 |
| | | ii | Item separation | 7.96 |
| | | iii | Person reliability | 0.63 |
| | | iv | Person separation | 2.08 |
| | | v | Internal consistency | 0.6 |

The first result which indicated the education policy makers found the SL-EdP comparatively easy is the person ability estimate mean of +0.77 logits. The items could be classified into four difficulty levels using the item measure (logit score), divided at the mean and standard deviation of the item or logit value of item (LVI) (Adams et al., 2020). With the item mean=0.00 and SD=1.43, the items were categorised into "very difficult", "difficult", "easy" and "very easy". As shown in Table 2, there were three items (15%) in the category of very difficult as answered by the policy makers (LVI > + 1.43 logit); seven items (35%) in the difficult category for item measure of +1.43 > LVI > 0; another eight items (40%) for the next category which is easy (0 > LVI > -1.43); and lastly, two items (25%) fell into very easy category (LVI < -1.43 logit).

*Table 2*
*Item Difficulty Level*

| | **Difficulty Level** | | | |
|---|---|---|---|---|
| | **Very Difficult** (LVI > +1.43) | **Difficult** (+1.43 ≥ LVI ≥ 0) | **Easy** (0 ≥ LVI ≥ -1.43) | **Very Easy** (LVI ≤ -1.43) |
| **Item** | SL12, SL3, SL15 | SL9, SL19, SL17, SL13, SL2, SL18, SL10 | SL7, SL14, SL8, SL16, SL20, SL11, SL5, SL6, | SL1, SL4 |
| **Number (%)** | 3 (15%) | 7 (35%) | 8 (40%) | 2 (10 %) |

Only SL1 and SL4 tended to be very easy for the education policy makers to answer, whilst most of the items, 15 out of 20 items (75%) were categorized as "difficult" and "easy" or could be considered as moderate questions. This indicated that policy makers do not have much difficulty to read, communicate and interpret statistical information.

***Distribution of Policy Makers Across Level***

The education policy makers could be categorised into four levels of statistical literacy namely very high level, high level, moderate and low level of statistical literacy. The logit value of person (LVP) was used to determine the level cut-off and the demographic characteristics profiling of policy makers in this study as tabulated in Table 3.

Approximately 17.4% of the policy makers are in very high level of statistical literacy while 14.02% other are in the high level. Majority of the respondents (59.15%) are of moderate level and only a few are at the low level with only 31 persons (9.5%). With regards to gender, at the highest level of statistical literacy, consists of 18.9% of the female respondents and only 14.8% of their male counterpart. While at the moderate level, the male percentage (65.6%) surpass the female (55.3%). For the high and low level, both genders had almost equal percentage.

In terms of age group, at the highest level of statistical literacy, showed that a higher percentage of older respondents belong to the high level. At the youngest age group, ranging from 21 to 30 years old, none were categorized in the very high level of statistical literacy, while their senior of 31-40 age group was about 14.5%, 41-50 at about 17.9% and the eldest age group of 51-60, 22.2% were included in the highest level of statistical literacy.

Analysis of the highest academic qualifications, PhD holders have the lowest percentage (12.4%) in the highest level of statistical literacy, while Master's degree holders have a better percentage (19.5%) and the ones with Bachelors' degree has the highest percentage (28.6%). Policy makers categorised in the high level of statistical literacy also consisted of mostly the Bachelors' degree holder (28.6%), followed by doctorate level (12.4%) and lastly Masters' degree holders (11.3%). Contrarily, in the moderate level, PhD holders has the highest

percentage, where more than half of them (64.1%) are in the moderate level followed by 60.9% of the Masters' holder and 35.7% from the Bachelors' degree.

In terms of grade, only one respondent is from the Jawatan Utama Sektor Awam (Jusa) or the highest grade in public sector is in the moderate level of statistical literacy. At the highest level of statistical literacy, the DG54 is the highest percentage, then, towards DG41, the lower the grade, the lower the percentage at the particular level. The trend was also seen in the high statistical literacy level. At the moderate level, Jusa dominates with the highest percentage, followed by DG54, DG44, DG52 and DG41. Whilst at the low statistical literacy level, the highest percentage was DG41, and percentage decreases as the grade increases except that none of the JUSA and DG54 officers were included in the low statistical literacy level.

*Table 3*
*Demographic Profiling of Education Policy Makers*

| | **Very High** | | **High** | | **Moderate** | | **Low** | | **Total** |
|---|---|---|---|---|---|---|---|---|---|
| | (LVP > + 1.77) | | (1.77 ≥LVP ≥ 1) | | (1 > LVP ≥ -0.23) | | (LVP < -0.23) | | |
| **Gender** | No | (%) | No | (%) | No | (%) | No | (%) | |
| **Male** | 18 | 14.8 | 14 | 11.5 | 80 | 65.6 | 10 | 8.2 | 122 |
| **Female** | 39 | 18.9 | 32 | 15.5 | 114 | 55.3 | 21 | 10.2 | 206 |
| | | | | | | | | | |
| **Age Group** | | | | | | | | | |
| **21-30** | 0 | 0.0 | 1 | 33.3 | 2 | 66.7 | 0 | 0.0 | 3 |
| **31-40** | 17 | 14.5 | 14 | 12.0 | 73 | 62.4 | 13 | 11.1 | 117 |
| **41-50** | 26 | 17.9 | 20 | 13.8 | 85 | 58.6 | 14 | 9.7 | 145 |
| **51-60** | 14 | 22.2 | 11 | 17.5 | 34 | 54.0 | 4 | 6.3 | 63 |
| | | | | | | | | | |
| **Highest academic qualification** | | | | | | | | | |
| **PhD** | 19 | 12.4 | 19 | 12.4 | 98 | 64.1 | 17 | 11.1 | 153 |
| **Master** | 26 | 19.5 | 15 | 11.3 | 81 | 60.9 | 11 | 8.3 | 133 |
| **Bachelor** | 12 | 28.6 | 12 | 28.6 | 15 | 35.7 | 3 | 7.1 | 42 |
| | | | | | | | | | |
| **Grade** | | | | | | | | | |
| **Jusa** | 0 | 0.0 | 0 | 0.0 | 1 | 100 | 0 | 0.0 | 1 |
| **DG 54** | 3 | 75.0 | 1 | 25.0 | 0 | 0.0 | 0 | 0.0 | 4 |
| **DG 52** | 9 | 26.5 | 7 | 20.6 | 16 | 47.1 | 2 | 5.9 | 34 |
| **DG 48** | 16 | 14.8 | 14 | 13.0 | 69 | 63.9 | 9 | 8.3 | 108 |
| **DG 44** | 27 | 15.9 | 22 | 12.9 | 103 | 60.6 | 18 | 10.6 | 170 |
| **DG 41** | 2 | 18.2 | 2 | 18.2 | 5 | 45.5 | 2 | 18.2 | 11 |
| | | | | | | | | | |
| **Tenure in service** | | | | | | | | | |
| **<10 years** | 1 | 8.3 | 4 | 33.3 | 6 | 50.0 | 1 | 8.3 | 12 |
| **10-19 years** | 22 | 14.6 | 18 | 11.9 | 93 | 61.6 | 18 | 11.9 | 151 |
| **20-29 years** | 28 | 22.0 | 15 | 11.8 | 74 | 58.3 | 10 | 7.9 | 127 |
| **30-39 years** | 6 | 15.8 | 9 | 23.7 | 21 | 55.3 | 2 | 5.3 | 38 |
| | | | | | | | | | |
| **Service institution** | | | | | | | | | |
| **MOE** | 38 | 20.0 | 35 | 18.4 | 118 | 58.4 | 6 | 3.2 | 190 |
| **SED** | 10 | 20.0 | 2 | 4.0 | 29 | 60.0 | 8 | 16.0 | 50 |
| **DOE** | 9 | 10.2 | 9 | 10.2 | 47 | 60.2 | 17 | 19.3 | 88 |
| | 57 | 17.4 | 46 | 14.0 | 194 | 59.1 | 31 | 9.5 | 328 |

This demographic profiling was also done in graphic form for easier digestion through visual comparison.

*Figure 1*
*Profiling of Education Policy Makers*

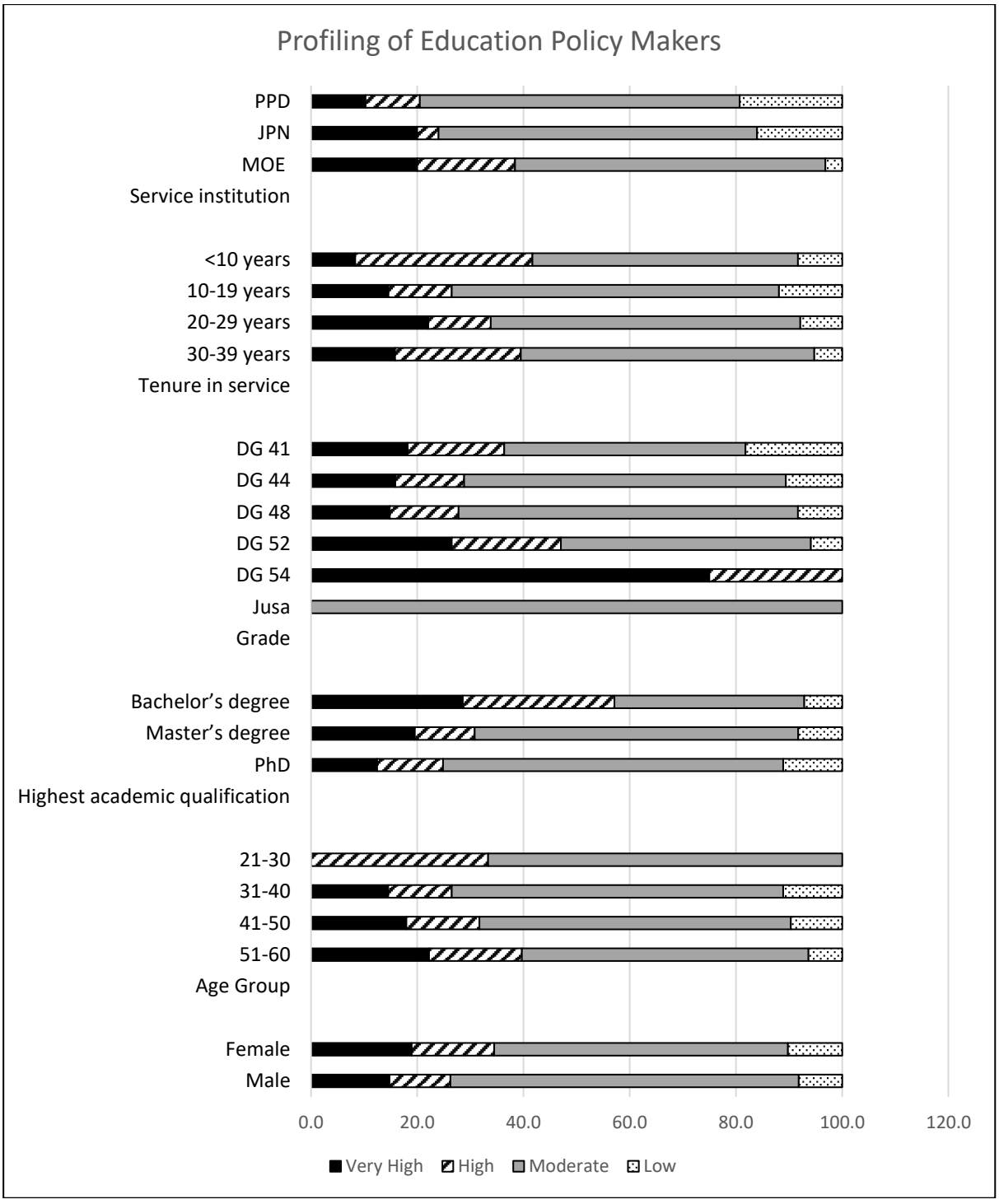


Figure 1 illustrated the same profiling in Table 4 in horizontal bar graphs for easier navigation and digestions. It is clearly seen here that education policy makers at the moderate level dominates in almost all demographic profiles.

***Differential Item Functioning (DIF) of Education Policy Makers' Demographic Characteristics in Statistical Literacy***

Next, the potential systematic differences in the scores of the item or item bias were examined across all the items. This analysis aimed not to have DIF as the items should be in similar level of difficulty disregard of any demographic characteristics. If items have a DIF between groups, that will imply that the items' difficulty level differs between certain demographics of the respondents, possibly suggesting the need to either reviewing the items or the scoring of the items should be scored differently. Instrument developers need to use certain quality control and statistical procedures to make sure that items in the instrument are fair for all respondents (Boone et al., 2014; Bond & Fox, 2015)

Demographic characteristics were analysed to identify if there are presence of DIF at the item level to explore whether the items in SL-EdP could measure and define statistical literacy in the same way for all categories of subpopulations. Table 4 shows summary of DIF criteria responses if DIF size > 0.5 logits, -2< t < 2.0 and p < 0.05. However, it is important to note that all values in the DIF criteria comparison should appear in all subgroups to confirm the presence of DIF. If the DIF criteria values only appear in either one subgroup, it does not indicate bias, but that the item is somehow more easily to comprehend by one group in the population compared to another (Boone et al., 2014). Furthermore, small size DIF can ignored in some cases, unless if an item is showing strong bias against another group, then that specific item should be dropped from the instrument. Table 4 illustrates the summary of DIF criteria.

*Table 4*
*Summary of DIF Criteria*

| | Criteria | Recommended Value | Authors |
|---|---|---|---|
| 1 | DIF size | >0.5 logits | Boone et al. (2014) |
| 2 | Rasch Welch t-test | -2 < t < 2.0 | Bond & Fox (2015) |
| 3 | Mantel Haenszel Probability | p < 0.05 | Boone et al. (2014) |

The DIF analysis showed that there are no DIF or bias items in this study as there were no demographic showing values in the DIF criteria in all subgroups to confirm the presence of DIF. However, there are three items, SL6, SL9 and SL15 which were DIF flagged, where only one subgroup in the demographic characteristics showed the values of a significant DIF indicating that the item is somehow are more difficult or easier to be answered by one subgroup as compared to another. Two demographic properties, which are tenure of service and service institution with one subgroup showing significant difference responses. However, the other four demographic factors involving gender, age, highest academic qualification, and grade did not have significant different responses among education policy makers. These DIF flagged items and the related subgroups were summarised in Table 5. These items were not dropped because they do not show bias but the ability of the particular subgroup in better understand the item were discussed.

*Table 5*
*DIF-flagged items in SL-EdP*

| Item Number | Demographic | Group | DIF size | t | p |
|---|---|---|---|---|---|
| **SL6** | Tenure of service | <10 years | -1.81 | -0.98 | 0.35 |
| | | 10-19 years | -2.4 | -1.02 | 0.31 |
| | | 20-29 years | 0.04 | 0.18 | 0.86 |
| | | **30-39 years** | **0.96** | **2.54** | **0.02** |
| **SL9** | Service institution | MOE | 0.18 | 1.17 | 0.24 |
| | | SED | **-0.67** | **-2.11** | **0.04** |
| | | DOE | -0.04 | -0.19 | 0.85 |
| **SL12** | Service institution | MOE | -0.18 | -1.06 | 0.29 |
| | | SED | -0.48 | -1.44 | 0.16 |
| | | DOE | **0.95** | **2.70** | **0.00** |

Table 6 further detailed out the items which have been found as DIF flagged to examine the reason as to why does a certain subgroup has superior or inferior difficulty in responding correctly compared to their counterparts.

*Table 6*
*Details of the DIF Flagged Items*

<table>
<tr><th></th><th>Item</th><th>Statement</th><th>Demographic with DIF</th></tr>
<tr><td>1</td><td>SL6</td><td>Table 1 shows times (in seconds) recorded for three students in seven 100-metre races.<br><table><tr><th>RACE</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th></tr><tr><td>Chan</td><td>14.2</td><td>13.8</td><td>14.0</td><td>13.7</td><td>13.3</td><td>13.5</td><td>13.5</td></tr><tr><td>Raju</td><td>14.8</td><td>14.7</td><td>14.4</td><td>14.8</td><td>13.8</td><td>13.6</td><td>13.5</td></tr><tr><td>Man</td><td>14.6</td><td>14.5</td><td>13.8</td><td>14.1</td><td>13.5</td><td>13.7</td><td>13.5</td></tr></table><br>Which student would you select for Majlis Sukan Sekolah Daerah (MSSD) the championships and why?</td><td>Tenure of service</td></tr>
<tr><td>2</td><td>SL9</td><td>In a newspaper article it was stated that in 50 years, in a city, the number of students had grown by 300%, from 30,000 to 90,000. Is this statement correct?</td><td>Service institution</td></tr>
<tr><td>3</td><td>SL15</td><td>The Report of Household Income and Basic Amenities Survey 2016 (Department of Statistics Malaysia) revealed the mean monthly income for Malaysia has increased RM6,141 in 2014 to RM6, 598 in 2016. Kelantan recorded the lowest mean monthly income of RM 4,214.<br><br>Based on these findings, if a student living in Kelantan with a family of monthly income of RM3,000, would she/he be considered for financial aid?</td><td>Service institution</td></tr>
</table>

Based on the responses, one item, SL6, showed a significant difference based on tenure of services. Policy makers who have been working for more than 30 years tend to be able to calculate the best average score for selection of students in MSSD better compared to their junior counterparts. This is probably due to the advantage of their longer working experience.

For service institution, two items, SL9 and SL15 showed DIF. In terms of correctly calculating percentage of student growth, the education officers in the state level SED found this difficult compared to officers in at federal and district level. For giving consideration for financial aid to give to students, the policy makers in DOE at district level did a better job and found it is easier compared to their counterparts in the state and federal level. Education officers working at the state level, State Education Department (SED) or better known as JPN found both items SL9 and SL15 difficult.

**DISCUSSION**

Generally, more than half (59.1%) of the policy makers were at the moderate level of statistical literacy, followed by 17.4% at the highest level, and 14% at the high level while low level of statistical literacy only consisted of 9.5% of the education personnel. The level of statistical literacy using cut-off logit value of person and Differential Item Functioning (DIF) was not conducted in previous statistical literacy studies as the main objectives were to assess on the items' hierarchy and quality only (Sabbag et al., 2018; Ziegler & Garfield, 2018).

This finding has however managed to bring the issues discussed in the problem statement where the key findings from the Survey of Adult Skills (OECD, 2013) that adults in almost all countries has low numeracy skills, this might not be the case for the education policy makers in Malaysia. Considering the officers' various background,

where some of them were language and religious teachers with minimum statistics background statistics since it was not made compulsory in their undergraduate studies (Reston et al., 2014) this finding showed that they were able to reach a moderate level of statistical literacy. This may be due to fact that most officers have encountered sufficient statistics in reports and documents to have the knowledge and skills to read, communicate and interpret statistical diagnoses and statements in their professional context.

Both genders scored almost the same at different level of statistical literacy. Personnel with Bachelors' degree scored better than master's degree and PhD holders. This may be due to that a bachelor's degree in fields associated with mathematics may score better in statistical literacy compared to a PhD in non-mathematical courses. Recommendation to investigate other demographic characteristics related to statistical literacy such as mathematics background in future studies. For age, working experience and service grade Supported the findings of a study by Pierce et al. (2013) where teachers with leadership positions demonstrate better statistical literacy as they have more reports to analyse.

The findings also showed that most education policy makers could successfully answer items in the levels of reading and comparing data which were represented by the inner circles in the framework for Professional Statistical Literacy (Pierce et al., 2014). However, only a small number of officers could correctly answer items focusing on analysing data. This is also consistent with previous literatures on school principals and teachers where Chick et al., (2014); Chick and Pierce (2013); Pierce and Chick (2013); Pierce et al., (2013; 2014). The study also concluded that the hierarchy of the items reflects the levels suggested in the professional statistical literacy framework and teachers showed no difficulty in reading data in straightforward presented graphs, charts and tables but appeared to have greater difficulty when needed deeper and higher-level analysis of data set, involving technical aspects of data presentation in the professional context (Pierce et al., 2014). Analysing is the highest level of challenge representing the highest level in the hierarchy framework, thus explaining the greater demand and requirement of professional context knowledge, thus making it more challenging for the respondents to get the right answer.

From the findings, interestingly, degree holders scored better compared to PhD holders. This could suggest more useful information on demographic characteristics such as mathematics or statistics background be included in future studies (Gönülal, 2018; Sesé et al.,2015). The study was done on education policy makers which was a homogeneous sample and could be expanded to other government personnel and policy makers. A more heterogeneous sample may provide more interesting results, thus, when the research respondents differ in their demographic background, the findings may provide clearer ideas on the impact of demographic characteristics on their statistical literacy. The empirical findings may affect item changes; thus, it would be interesting to explore whether the findings of this study could be replicated in such studies.

## CONCLUSION

Profiling of the policy makers provided insights on area of strength and weaknesses within level of statistical literacy of personnel based on their demographic characteristics. This would make SL-EdP as a useful instrument in evaluating statistical literacy for many purposes such as a self-reflection tool for statistical training needs and professional development program planned for education policy makers. Moreover, this profile pertaining statistical literacy could be recommended for recruitment and placement of personnel as well as suggestions for interventions. The findings provided information on which topics, skills, and aspects in statistical literacy that the policy makers are struggling with, and which parts do not need further attention. Educators and instructors of professional development program might use the findings to help develop training curriculum as well as evaluate the effectiveness of such programs and further improvise future interventions.

## REFERENCES

Adams, D., Joo, M. T. H., Sumintono, B., & Pei, O. S. (2020). Blended learning engagement in higher education institutions: A differential item functioning analysis of students' backgrounds. *Malaysian Journal of Learning and Instruction*, *17*(1), 133–158.

Akanmu, S., & Jamaludin, Z. (2015). A conceptual framework for students' data-driven decision making process in higher education institutions. *Advanced Science Letters*, *21*(7), 2256–2260. https://doi.org/10.1166/asl.2015.6275

Bond, T., & Fox, C. (2015). *Applying the Rasch model : Fundamental measurement in the human sciences* (3rd ed.). Routledge Taylor & Francis Group.

Boone, W. J., & Noltemeyer, A. (2015). Rasch analysis: A primer for school psychology researchers and practitioners. *Cogent Education*, *4*(1), 1–13. https://doi.org/10.1080/2331186X.2017.1416898

Budgett, S., & Rose, D. (2017). Developing statistical literacy in the final school year. *Statistics Education Research Journal*, *16*(1), 139–162.

Callingham, R., & Watson, J. (2017). The development of statistical literacy at school. *Statistics Education Research Journal*, *16*(1), 181–201.

Chick, H., & Pierce, R. (2012). The statistical literacy needed to interpret school assessment data. *Mathematics Teacher Education & Development*, *15*(2), 1–19. https://login.ezproxy.utas.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=94988608&site=eds-live

Chick, H., Pierce, R., & Wander, R. (2014). Sufficiently assessing teachers' statitistical literacy. *Proceedings of the Ninth International Conference on Teaching Statistics*, 1–6.

Gal, I. (2002). Adults' statistical literacy : Meanings , components , responsibilities. *International Statistical Review*, *70*(1), 1–51.

Gal, I., & Ograjensek, I. (2017). Official statistics and statistics education : Bridging the gap. *Journal of Official Statistics*, *33*(1), 79–100.

Gönülal, T. (2018). An investigation of the predictors of statistical literacy in second language acquisition. *Eurasian Journal of Applied Linguistics*, *4*(1), 49–70. https://doi.org/10.32601/ejal.460631

Hanuschek, E., & Woessmann, L. (2020). *The economic impacts of learning losses* (Issue September). OECD. https://www.oecd.org/education/The-economic-impacts-of-coronavirus-covid-19-learning-losses.pdf%0Ahttps://www.oecd-ilibrary.org/education/the-economic-impacts-of-learning-losses_21908d74-en

Henderson, J., & Corry, M. (2020). Data literacy training and use for educational professionals. *Journal of Research in Innovative Teaching & Learning*, *ahead-of-p*(ahead-of-print). https://doi.org/10.1108/jrit-11-2019-0074

Kaplan, J. J., & Thorpe, J. (2010). Post secondary and adult statistical literacy: Assessing beyond the classroom. *Eighth International Conference on Teaching Statistics*.

Krishnan, S. (2014). *Assessing students' statistical literacy of inferential statistics using Rasch analysis*. (Unpublished Doctoral Thesis).

Maki, S., & Horita, T. (2018). Comparative study of the categorization of items of statistical literacy in Mathematics textbooks of elementary , junior high , and high schools in Japan. *International Journal of Learning Technologies and Learning Environments*, *1*(1), 79–92.

Ministry of Education. (2013). Malaysia Education Blueprint 2013 - 2025. In *Ministry of Education* (Vol. 27, Issue 1). https://doi.org/10.1016/j.tate.2010.08.007

Noor Lide, A. K., Nor Zatul-Iffa, I., Zamalia, M., & Mohammad Said, Z. (2010). Measuring students' understanding of statistical concepts using Rasch measurement. *International Journal of Innovation, Management and Technology*, *1*(1), 13. https://doi.org/http://dx.doi.org/10.7763/IJIMT.2010.V1.3

OECD. (2013). *Skilled for Life ? Key Findings from the Survey of Adult Skills*. OECD.

Piaget, J. (1972). *Psychology and Epistemology: Towards a Theory of Knowledge*. Grossman Pierce, R., & Chick, H. (2013). Workplace statistical literacy for teachers: Interpreting box plots. *Mathematics Education Research Journal*, *25*(2), 189–205. https://doi.org/10.1007/s13394-012-0046-3

Pierce, R., Chick, H., & Gordon, I. (2013). Teachers ' perceptions of the factors influencing their engagement with statistical reports on student achievement data. *Australian Journal of Education*, *57*(3), 237–255. https://doi.org/10.1177/0004944113496176

Pierce, R., Chick, H., Watson, J., Les, M., & Dalton, M. (2014). A statistical literacy hierarchy for interpreting educational system data. *Australian Journal of Education*, *58*(2), 195–217. https://doi.org/10.1177/0004944114530067

Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, *15*(2), 20111. https://doi.org/10.1103/PhysRevPhysEducRes.15.020111

Reeves, T. D., & Chiang, J. L. (2018). Online interventions to promote teacher data-driven decision making: Optimizing design to maximize impact. *Studies in Educational Evaluation*, *59*(August), 256–269. https://doi.org/10.1016/j.stueduc.2018.09.006

Reston, E. (2010). Statistical literacy assessment and training of government personnel using data from national statistics office: Philippine context. *Eighth International Conference on Teaching Statistics*.

Reston, E., Krishnan, S., & Idris, N. (2014). Statistics education research in Malaysia and the Philippines: A comparative analysis. *Statistics Education Research Journal*, *13*(2), 218–231.

Sabbag, A., Garfield, J., & Zieffler, A. (2018). Assessing statistical literacy and statistical reasoning. *Statistics Education Research Journal*, *17*(2), 141–160. http://www.stat.auckland.ac.nz/serj%0Ahttps://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_8C2.pdf

Schield, M. (2016). Augsburg student evaluations of Stat 102: Social statistics for decision makers. *2016 ASA Proceedings of the Section on Statistical Education*, *1*(1), 3741–3750.

Sesé, A., Jiménez, R., Montaño, J. J., & Palmer, A. (2015). Can attitudes toward statistics and statistics anxiety explain students' performance? *Journal of Psychodidactics*, *20*(2), 285–304. https://doi.org/10.1387/revpsicodidact.13080

Sharma, S. (2017). Definitions and models of statistical literacy: A literature review. *Open Review of Educational Research*, *4*(1), 118–133. https://doi.org/10.1080/23265507.2017.1354313

Sireci, S. G., & Rios, J. A. (2013). Decisions that make a difference in detecting differential item functioning. *Educational Research and Evaluation*, *19*(2–3), 170–187. https://doi.org/10.1080/13803611.2013.767621

Tiro, M. A., Aidid, M. K., & Ahmar, A. S. (2018). Exploration of table and graph literacy of statistics student at Universitas Negeri Makassar. *Journal of Physics*, *1040*(1), 1–5.

UNESCO. (2018). *Handbook on Measuring Equity in Education*. UNESCO Institute for Statistics.

United Nations Economic Comission for Europe. (2012). Making data meaningful : A guide to improving statistical literacy. In *United Nations Economic Comission for Europe: Vol. Part 4*. https://doi.org/10.1111/j.1600-0897.1998.tb00338.x

Wallman, K. K. (1993). Enhancing statistical literacy: enriching our society. *Journal of the American Statistical Association*, *88*(421), 1–7. https://doi.org/10.2307/2290686

Watson, J., & Callingham, R. (2020). COVID-19 and the need for statistical literacy. *Australian Mathematics Education Journal*, *2*(2), 20–25.

Wright, B., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement*. MESA Press.

Yolcu, A. (2014). Middle school students' statistical literacy: Role of grade level and gender. *Statistics Education Research Journal*, *13*(2), 118–131.

Ziegler, L., & Garfield, J. (2018). Developing a statistical literacy assessment for the modern introductory statistics course. *Statistics Education Research Journal*, *17*(2), 161–178.