©2016 Journal of Contemporary Islamic Studies (JCIS) is a fully refereed journal published annually by the Academy of Contemporary Islamic Studies, Universiti Teknologi MARA Shah Alam Malaysia, 40450 Shah Alam Selangor, Malaysia

Disclaimer:
The views and opinions expressed therein are those of the individual authors, and the publication of those statements in JCIS does not imply endorsement by the publisher or the editorial staff. Copyright is reserved jointly by Universiti Teknologi MARA and Academy of Contemporary Islamic Studies (ACIS).

The illustrated image at the front cover of JCIS is a diwan's calligraphic art of Qur'anic verse: وَقُولُ ۗ رَبِّ زِبَرُ عَلَيْكَ which means "...and say, "my Lord! increase me in knowledge" (Thoha: 114). The verse is indeed, implies our prayer to call for the enhancement of our knowledge. The true knowledge that can lead to the absolute truth and prosperity; combination of revealed and contemporary knowledge. Therefore, JCIS is one of the medium for disseminating these knowledge.
CONTENTS

The Reasoning Pattern of Islamic Jurists' Views on Ar-Rahn (Islamic Pawn Broking) Contract and Its Ruling
Dziauddin Sharif, Amir Shaharuddin & Nurul Aini Muhamed

Application of Rasch Measurement Model in Quality Evaluation of Shariah Compliant Gold Investment (SCGI) Instrument
Najahudin Lateh & Ghafarullahhuddin Din

The Best Interest of the Child: An Analysis of the Judicial Approaches in Assessing Child Maintenance Applications in Shariah Courts
Nurhidayah Hashim & Tim Lindsey

An Overview of Ibra' Implementation in Malaysian Islamic Banks
Mohd Ab Malek Md Shah, Jeniwayt Mohd Jody, Mohd Harun Shahudin & Sulaiman Mahzan

Behavioural Response Patterns on the Organisational Factors Influencing Compliance Behaviour of Business Zakat
Mohd Rahim Khamis

Interactive Multimedia for Learning Hajj
Ilinadia Jamil & Zainal Kadir

JOURNAL OF CONTEMPORARY ISLAMIC STUDIES
Kesedaran Pelabur Harta Tanah Terhadap Sistem Kewangan Islam
Rohayu Ab. Majid & Rosli Said

Kekerapan Penerbitan di SCOPUS Mengenai Flora-Flora dalam Al-Quran dan Al-Hadith
Mohd Asmadi Yakob, Mohd Yakub @ Zulkifli Mohd Yusoff, Monika @ Munirah Abd Razzak Khadher Ahmad, Nurulwahidah Fauzi, Khalijah Awang, Rozana Othman & Mohd Rais Mustafa

Perkongsian Pendapatan Suami Isteri dalam Keluarga Islam dan Impak kepada Kualiti Hidup
Azhar Abdul Aziz, Raihanah Hj Azahari & Khairul Faezi Mohd Said

تطور الاستشراق
Haitham Ibrahim Ahmad al-Sharu, Salmah Ahmad & Muhammad Nur Abdullah
Application of Rasch Measurement Model in Quality Evaluation of Shariah Compliant Gold Investment (SCGI) Instrument

Najahudin Lateh* & Ghafarullahhuddin Din

1,2 Academy of Contemporary Islamic Studies (ACIS)
Universiti Teknologi MARA (UiTM),
40450 Shah Alam, Malaysia

*Corresponding Author
najahudin@salam.uitm.edu.my

ABSTRACT
This paper describes the process of assessing the validity and reliability of a new instrument namely Shariah Compliant Gold Investment (SCGI). The instrument consists of 33 items that are embedded in three dimensions and was administered to 27 Malaysian investors and investment institutions. Rasch model was used to examine the validity of items by two criteria (1) point measure correlation (PTMEA CORR) and (2) fit statistics (infit/outfit MNSQ and z-std). The findings indicated that the reliability value for the respondents and items are high with r = 0.91 and r = 0.81 respectively with Cronbach alpha 0.93. At the same time, the item separation is 2.07 while the person separation is 3.15. As for the items polarity, most of the items contributed to the measurement as all of the PTMEA CORR values are positive (+0.44 logit to +1.66 logit) except for the A03 item (0.17 logit). The fit item testing indicated that the value of the sum of the mean of infit MNSQ and SD was between +0.68 logit to +1.30 logit. Only one item (A03) falls in the range of elimination due to negative value of PTMEA CORR and z-std > 2.0. The results suggested the item to be removed, retaining the balance of 32 items.

Keywords: Gold Investment, Rasch Model, Shariah Compliant

1. Introduction

A shariah-compliant gold transaction has been authentically justified in a few hadith, among them narrated by ‘Ubadah ibn al-Samit in which the Prophet Muhammad SAW said: “Gold (exchanged) with gold, silver for silver, wheat for wheat, barley for barley, salt for salt, and they should be
of equal weight, and hand to hand. If the types of goods exchanged are different, then sell without delay and submit the goods directly.” (Muslim, 2010). The hadith pointed two conditions for a shariah-compliant gold investment; equal and on-the-spot transaction (al-Sharbini, 1978; al-Saddam, 2006).

In Malaysia, both criteria have been gazetted as Gold Investment Parameter, endorsed by the National Fatwa Council. It functions as guidance for investors as well as investment institutions. However, the parameters are too general. This has urged the Shariah Advisory Council of Malaysia to call for the parameters to be reviewed (Jakim, 2012). Recently, Najahudin et al. (2014) propose Shariah Compliant Gold Investment (SCGI) as a new guideline. Thus, this research aims to evaluate the validity and reliability of the SCGI using Rasch Measurement Model. The Rasch analyses will be focused on the interpretation of data reliability, item polarity, fit statistics and the persons-items distribution map.

2. Shariah Compliant Gold Investment (SCGI)

The SCGI has been developed meticulously through systematic procedures involving relevant experts (Najahudin et al., 2014). It is more specific and consists of three dimensions; (i) investor and investment institutions; (ii) products and prices; and (iii) the contracts offered. These three dimensions and a total of 33 items have been unanimously agreed by 13 experts via two rounds of Delphi technique. Each round was implemented using a questionnaire with 4-Point Likert Scale; (1) strongly disagree, (2) disagree, (3) agree, and (4) strongly agree. Data collected from each round were analyzed using Statistical Package for Social Science (SPSS) in order to attain the agreed dimensions and items. In the second round, the expert consensus had successfully obtained. All items indicated that the consensus were in the interquartile range of (IQR) = 0 to 1, median = 4 and mode = 4, above the 95 percentage. The median frequency distribution of this study is consistent with Green’s (1981) value of 3.8. The items which have been agreed upon are shown in Table 1.
Table 1. Dimension and Items of SCGI

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Number of Items</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investor and investment institutions</td>
<td>A01, A02, A03, A04, A05</td>
<td>5 items</td>
</tr>
<tr>
<td>Product and prices</td>
<td>B01, B02, B03, B04, B05, B06, B07, B08, B09, B10, B11, B12, B13, B14, B15, B16, B17</td>
<td>17 items</td>
</tr>
<tr>
<td>Contract deal</td>
<td>C01, C02, C03, C04, C05, C06, C07, C08, C09, C10, C11</td>
<td>11 items</td>
</tr>
</tbody>
</table>

3. Methodology

Source of Data
For the purpose of validating 33 items of SCGI, the researcher organized a special seminar on 4 April 2015. The seminar attracted 27 participants who were of gold investors. Prior to administering the SCGI, the researcher thoroughly explained the dimensions and items to ensure the respondents' understanding correspond to the questions. At the end of the seminar, 27 valid responses were collected.

3.1 Rasch Measurement Model

The Rasch model is a measurement on the probability of interaction between the person and the item. Each person will be categorized based on their temporary skills whereas the items are categorized based on their difficulty. The Rasch model was formed by taking into consideration the ability of the person answering the questionnaire or the instrument and the difficulty posed by each of the question or the item. The ability of the person and the difficulty of the item were shown in the form of logit through the transformation of ordinal data into ratio measurements. This model would be able to predict the pattern of the response based on the different ability of each person and the difficulty of each item (Rasch, 1980). The probability to succeed would depend on the difference of the ability of the respondent and the difficulty of the item. According to Rasch, (i) a smarter person would have a bigger probability to agree with the items; and (ii) items that are less difficult would have a higher probability to be agreed by all of the respondents (Bond & Fox, 2015).

The Rasch model is able to provide the accuracy of the validity and reliability as it focuses on the person and the item. Moreover, this model
would be able to show which of the item or construct would fit, misfit, requires further research or eliminated (Azrilah, 2010) based on the established rating scale. This study utilised the acceptable ranges of the Rasch model as shown in Table 2.

Table 2. Rating Scale Instrument Quality Evaluation Criteria using Rasch Model

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Statistical Info</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item Validity</td>
<td>a. Item Polarity PTMEA CORR &gt; 0.4 – 0.8 (Linacre, 2011; Azrilah, 2010)</td>
<td>PTMEA CORR &gt; 0.4 – 0.8 (Linacre, 2011; Azrilah, 2010)</td>
</tr>
<tr>
<td>Item Item Misfit</td>
<td>b. Item Fit Total MNSQ infit and outfit of 0.5 - 1.5 (Linacre, 2011; Linacre, 2002)</td>
<td>Total MNSQ infit and outfit of 0.5 - 1.5 (Linacre, 2011; Linacre, 2002)</td>
</tr>
<tr>
<td>Item Misfit</td>
<td>c. Separation All items show ≥ 2.0 (Linacre, 2011; Fisher, 2007)</td>
<td>All items show ≥ 2.0 (Linacre, 2011; Fisher, 2007)</td>
</tr>
<tr>
<td></td>
<td>d. Person Reliability Value &gt; 0.8 (Bond &amp; Fox, 2015)</td>
<td>Value &gt; 0.8 (Bond &amp; Fox, 2015)</td>
</tr>
<tr>
<td></td>
<td>e. Item Reliability Value &gt; 0.8 (Bond &amp; Fox, 2015)</td>
<td>Value &gt; 0.8 (Bond &amp; Fox, 2015)</td>
</tr>
</tbody>
</table>

Source: Bond & Fox (2015); Linacre (2011); Azrilah (2010); Fisher (2007); Linacre (2002); Wright & Stone (1979)

4. Data analysis

Data were analyzed using the Rasch analysis software, the WINSTEPS 3.72.3. Rasch predicts the probability of a person to evaluate item, and the probability for each item to be evaluated by a person. In Rasch Measurement Model, the validity of the instrument could be identified through several major analysis such as the item polarity, person-item fit, person-item misfit, the person-items distribution map, person-item separation, unidimensionality and scale calibration (Rasch, 1980; Bond & Fox, 2015; Linacre, 2011). This study only reports on the reliability value, item polarity, fit statistics and person-items distribution map (PIDM). Figure 1 summarises the types of analysis performed.
Summary Statistics:
- Realibility (Person and item)
- Cronbach alpha
- Separation (Person and item)

Fit Statistics:
- PTMEA CORR
- infit/outfit MNSQ
- infit/outfit z-std (Person fit, item fit, misfit responded and misfit items)

PIDM:
- Mean, maximum, minimum logit (person and item)
- Separation(person / item)
- Logit scale ruler

Dimension A: Investor and investment institutions (5 items)
Dimension B: Product and prices (17 items)
Dimension C: Contract deal (11 items)

Shariah-Compliant Gold Investment (SCGI) Instrument

Figure 1. Analysis and validation process
5. Results and discussion

5.1 Reliability
Reliability is the value that indicates the consistency of the position of the person and item in the logit scale. The person reliability value shows the consistency of the position of the respondent when given another set of items that measures the same construct. The item reliability value shows the consistency of the set of items when answered by different respondents who have similar abilities. The coefficient value that is closest to 1.00 denotes a high reliability (Nunnally & Bernstein, 1994).

According to Bond and Fox (2015) and Linarce (2011), the reliability of a person which exceeds 0.80 values indicates a strong acceptance towards the respondent or the item. On the other hand, Fisher (2007) divided the rating scale for the reliability of a person and item into poor (< 0.67), fair (0.67 - 0.80), good (0.81 - 0.90), very good (0.91 - 0.94) and excellent (> 0.94). The accepted separation value for a person and item must be at least 2.0 (Linarce, 2011; Fisher, 2007).

Based on Figure 2, the summary statistic displays acceptable person and item reliability values. On top of that, the Cronbach-α of 0.93 is good, indicating the instrument is a valid measurement and capable of identifying the level of shariah-compliance of gold investment products. The reliability of the item recorded a value of 0.81, which indicates that there are sufficient items to measure what need to be measured (Nunnally & Bernstein, 1994).

The respondent reliability value is 0.91, indicating a strong probability of the items to measure the same result when given to another similar respondent (Azrilah, 2010). In addition, the separation value for respondent and item were 3.15 and 2.07, respectively. A value of ≥ 2.0 is good, indicating the SCGI ability to segregate respondent ability and item difficulty.
5.2 Polarity of the Item

The item polarity is a precondition that must be referred to by reviewing the point measure correlation (PTMEA CORR) coefficient. Items are assumed to be able to differentiate the ability of the respondents when the PTMEA CORR values are high. The value must be positive to indicate the item is moving in parallel (Bond & Fox, 2015). The negative or zero PTMEA CORR values indicates that the response of a person or item conflicts with the variables constructed (Linacre, 2011), an inverse direction of measurement and an uncommon decision making variable (Azrilah, 2010). Nunnally and Bernstein (1994) and Finlayson (2009) believed that the PTMEA CORR item value of at least +0.30 logit would be able to measure a construct systematically, whereas a value of +0.32 logit would be able to merely measure in an average manner. However, this study uses the value between +0.4 logit and +0.8 logit (0.4 < x < 0.8) to prove that the constructed items would be measureable and able to differentiate the respondents (Linacre, 2011; Fisher, 2007; Azrilah, 2010).
Figure 3 shows that all the items have positive PTMEA CORR values and small mean error measurement of SE (+0.39 logit), except item A03, which reported a negative value of -0.17 logit (SE = +0.76 logit). This particular item was eliminated, as it did not measure what should be measured (Azrilah, 2010). Most of the values of the other items are between the values of +0.42 logit to +0.77 logit, except for 2 items that are outside the specified range that is A01 (+0.77 logit) and A02 (+0.15 logit). However, both items were retained, based on their acceptable infit MNSQ (+1.49 logit and +1.48 logit respectively) and z-std (1.7 and 0.6 respectively).

![Table 10.1](image-url)
5.3 Fit Statistics

The Rasch model provides fit statistics to detect item or person misfit. The fit statistics refer to (i) infit and outfit mean square (MNSQ); and (ii) infit and outfit standardized (z-std); for both person and items. MNSQ is the ratio of an observation compared to the expectation. The ideal value for MNSQ is 1, when the observation corresponds with the expectation. The MNSQ value is excluded from the expectation when the total mean value of the MNSQ infit and the SD (mean iMNSQ +/- SD) is out of the specified range.

According to Bond and Fox (2015), the values of the MNSQ infit and outfit for each person and items for the likert scale must be between +0.6 logit to +1.4 logit. Fisher (2007) established that the fit item has a fair scale of within +0.34 logit to +2.9 logit, whereas a good scale has a value of within +0.50 logit to +2.0 logit. However, this study utilised the range of values recognised by Linacre (2002) in which the values between +0.5 logit to +1.5 logit in order to verify the fit and misfit for a person or an item. Usually the outfit would be more sensitive to the response compared to the infit (Linacre, 2002). The detection of the items that are misfit or outlier can be further confirmed with the z-std values that must be between the ranges of -2.0 to 2.00. The ideal value for z-std is 1.0 (Azrilah, 2010). The person or item that does not fulfil the criteria range will be considered to be eliminated, except if the PTMEA CORR values for the person and item is between +0.4 logit and +0.8 logit.

This study focused on the fit item compared to the fit person. The fit item here means that the given value has an item function and is able to measure the latent trait required. Misfit occurs when (i) the item does not measure the desired traits; (ii) the items are too difficult or too simple for the person; (iii) or there was an unstable response from the person. Figure 3 shows the sum of the MNSQ infit mean and (+/-) SD (0.99 logit + (--) 0.31 logit) are among the values between +0.68 logit to +1.30 logit, which is acceptable.

All the items were accepted. However, item A03 was eliminated as it was outside the acceptable range of z-std (outfit 2.90) and has a negative PTMEA CORR value (-0.17 logit). Even though the infit MNSQ of item B14 (+1.66 logit), C06 (+1.51 logit) and B09 (+1.51 logit) were beyond the acceptable range (0.5 < y < 1.5), all of them were accepted as their z-std outfit were within the acceptable range (B14 = 1.8; C06 = 1.7; B09 = 1.4). They were also measuring in the right direction as the PTMEA CORR values were positive (B14 = +0.42; C06 = +0.56; B09 =
+0.55). Therefore, all of the items (n = 33) were retained except item A03.

5.4 Persons-Items Distribution Map (PIDM)

PIDM is the heart of the Rasch model analysis, which shows the hierarchical relationship of the ability of the person and the difficulty of the item (Bond & Fox, 2015). The person with a higher ability and a more difficult item is placed at the top, whereas a person with a lower ability and an easy item is placed at the bottom. Based on Figure 2, the mean value for the evaluation of a person is +2.77 logit and for the item is 0.00 logit. The minimum value for a person is +0.23 logit whereas the maximum value is +5.08 logit. The minimum value for the item is -3.19 logit whereas its maximum value is +2.90 logit. This makes the total ruler length of a person to be 5.31 logit against the item value of 6.09 logit. The gap that is lacking between the scale of the person compared to the measured item is 0.78 logit (6.09 - 5.31). This hierarchical value is shown in Figure 4.
Figure 4. Hierarchy of relationship
The PIDM above shows the ability of the item to separate the respondents into three categories namely the person free item, the person above the mean and the person below the mean. The items were divided into difficult and easy, with the item mean (0.00 logit) as the separation line. This division is aligned with the data separation of the person (3.15) and the item (2.07) as shown in Table 3. Out of the 27 respondents, Group 1 (excellent) contains 12 people which were located within the maximum item location at a range of values between +2.90 logit to +5.08 logit. Group 2 (good) consisting of 7 people were located within +2.90 logit to +2.07 logit. The rest of the respondents were in Group 3 (mediocre) as they were within the range of +2.07 logit to item mean (0.00 logit).

The map proves most of the respondents were person free item. More items are required to measure them. The respondents have high evaluation ratings, and they had no problem to agree with most of the items in the instrument. The items only measure the person in Group 2 and 3, whereas there was no complicated item to measure the people in Group 1. Most of the items are easy and below the respondent mean (+2.77 logit). There were no respondents under the mean item (0.00 logit). This is aligned with the view of Bond and Fox (2015) in which, an easier item is more likely to be agreed upon by all of the respondents.

The easiest item to be agreed by most of the respondents was the A03 (-3.19 logit) and the most difficult item to be agreed together was the A05 (+2.90 logit). There are also large gaps in two places namely between the items A01 and A05, and also the items B01 and A03. This made the item reliability to be at a value of 0.81.

5.5 Summary Statistic after Removal of Misfit Item

The value for the statistics analysis after the elimination of A03 item is shown in Table 3. Overall, the findings showed that the instrument has a fair item reliability (+0.75 logit), mean infit MNSQ (+1.02 logit) and mean outfit z-std (0.00).
6. Conclusion

The analysis of the Rasch model has proven that SCGI can be accepted and has a high reliability (person = 91; item = 81). Out of the original 33 items, 1 of them was a misfit and was required to be eliminated in order to obtain a valid instrument under the Rasch model. Therefore, the final instrument contained only 32 items. SCGI instrument can be used as guidance to the public and adopted by Islamic financial institutions to create new products or audit of the existing Shariah gold investment. Further studies can be carried out by extending the scope of management, public understanding, product selection factors and fraud in a gold investment. The study can also be made on the correlation and significant differences between the parties involved in gold investment such as investors and investment institutions or between banks and investment companies. The researchers then are able to use or modify the approach and the methodology of this study to develop other shariah-compliant muamalat instruments such as real estate investments or forex. In addition, the increase in dimension of shariah-compliant can also be done from the maqasid al-shariah, legal documents, financial reports or corporate governance.
7. References


Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean?. Rasch measurement transactions, 16 (2), 878.


