

## RESEARCH ARTICLE

## Factorial validity, measurement and structure invariance of the Malay language decisional balance scale in exercise across gender

Garry Kuan<sup>1,2\*</sup>, Abdulwali Sabo<sup>3</sup>, Sukanlaya Sawang<sup>4</sup>, Yee Cheng Kueh<sup>3\*</sup>

**1** Exercise and Sports Science, School of Health Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia, **2** Department of Life Sciences, Brunel University London, Uxbridge, United Kingdom, **3** Unit of Biostatistics and Research Methodology, School of Medical Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia, **4** International Centre for Transformational Entrepreneurship, Coventry University, Coventry, United Kingdom

\* [garry.kuan@brunel.ac.uk](mailto:garry.kuan@brunel.ac.uk) (GK); [yckueh@usm.my](mailto:yckueh@usm.my) (YCK)



## OPEN ACCESS

**Citation:** Kuan G, Sabo A, Sawang S, Kueh YC (2020) Factorial validity, measurement and structure invariance of the Malay language decisional balance scale in exercise across gender. PLoS ONE 15(3): e0230644. <https://doi.org/10.1371/journal.pone.0230644>

**Editor:** César Leal-Costa, Murcia University, SPAIN

**Received:** November 14, 2019

**Accepted:** March 4, 2020

**Published:** March 18, 2020

**Copyright:** © 2020 Kuan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** This research was supported by the University Research Individual (1001.PPSP.8012250) from Universiti Sains Malaysia (<http://www.research.usm.my>). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

## Background

This study examined the psychometric properties of the Malay version of the decisional balance (DB-M) for exercise (i.e. perceived benefits and perceived barriers) using a cross-sectional design. Also, this study assessed the measurement and structural invariance of the DB-M across gender.

## Methods

The study sample consisted of 750 students (female: 51.7%, male: 48.3%), with a mean age of 20.2 years (SD = 1.2). Decision balance (DB) scale was assessed with the 10-item DB-M. Standard forward-backward translation was performed to translate the English version of the DB into Malay version (DB-M).

## Results

The confirmatory factor analysis (CFA) results based on the hypothesised measurement model of two factors and ten items demonstrated adequate factor structure after the addition of some correlated item residuals (comparative fit index (CFI) = .979, Tucker and Lewis index (TLI) = .969, standardised root mean square residual (SRMR) = .037, root mean square error of approximation (RMSEA) = .047). The construct reliability and average variance extracted values were .850 and .839, and .542 and .538, for perceived benefits and perceived barriers, respectively. Meanwhile, the Cronbach's alpha was .857 and .859, and the intraclass correlation coefficient for test-retest reliability was .979 and .960 for perceived benefits and perceived barriers respectively. The findings provided evidence for measurement invariance of DB-M for the male and female samples. The final CFA model fit the data well for both male sample (CFI = .975, TLI = .964, SRMR = .040, RMSEA = .052) and female sample (CFI = .965, TLI = .949, SRMR = .044, RMSEA = .058).

## Conclusions

The translated version of the DB-M was valid and reliable for assessing the level of perceived benefits and perceived barriers in exercise among university students in Malaysia.

## Introduction

It has been widely reported that participation in regular physical activity can significantly reduce the risk of hypertension, diabetes, cardiovascular diseases, obesity, overall fitness, mental wellness, and quality of life [1, 2]. Frequent physical activity is an essential determinant that enhances the human body system (e.g. musculoskeletal, cardiovascular, and metabolic system) for optimal functioning [3]. However, in 2009, physical inactivity was reported as the fourth leading risk factor for noncommunicable diseases, causing about three million premature deaths [4]. In Malaysia, the prevalence of physical inactivity was 39.7% and that the population spent most of their time (74% of the day) in sedentary activities, such as sleeping or lying down [5].

The Transtheoretical Model (TTM) is among the most commonly used theories of the behaviour of change that has been successfully employed by researchers to describe a spectrum of health behaviours such as physical activity and health promotion [6]. The TTM is a collective psychological framework that aims to demonstrate deliberate health behaviour choice and sustenance as a process that transpires over time as a function of behavioural history and motivation [7]. This model illustrates how people move dynamically through stages of behavioural change [8]. Many studies have used the TTM or Stages of Behavioural Change (SBC) Model to develop interventions to promote physical activity [9, 10].

The TTM compromises the temporal dimension and the stages of change, and these were integrated into the process and principles of change of various testable theories. The TTM is applied to exercise behaviour to illustrate the capability and the extent of willingness and incremental shift over time, as well as its specification of transformation-related to and profile variations between phases [8]. The TTM, which depicts human readiness to change and advancement through a sequence of phases, involves five key constructs, i.e. stages of change, the process of change, decisional balance, self-efficacy and temptation [8]. This study focuses on the testing of the translated Malay version of the decisional balance components of the TTM.

The decisional balance construct reflects the struggle model, which is an essential process of making a decision associated with specific health practices [11]. Decisional balance signifies the perceived benefits (pros) and perceived barriers (cons) related to exercise behaviour. Perceived benefits for exercise refer to enhanced self-confidence, physical strength, and aerobic ability, while perceived barriers to exercise refer to physical distress, cost, and hesitation. Many researchers reported that the level of perceived benefits increases while the barriers decrease within the stages of behavioural change [8]. The decisional balance sheet (DBS) developed by Janis and Mann [11] is among the most assuring intervention for exercise domain [12]. The DBS is derived from the expectancy theory, which stated that the level of a person's expectations of gains or failures is what determines his relative course of action. The more critical the information is viewed, before reaching a decision, the more effective the commitment to that decision and the more stable the adherence to the decision [11]. A study by Nigg and Courneya [12] confirmed that the DBS intervention was effective in maintaining current exercise

behaviours, whereas, a decline in exercise behaviours was witnessed among the placebo groups over eight weeks.

The Decisional Balance (DB) scale was initially developed by Marcus and Rakowski [13], consisting of two subscales: (1) perceived benefits (pros, ten items) reflect the positive aspects of exercise and (2) the perceived barriers (cons, six items) reflect the negative aspects of exercise. The decisional balance scale was later revised and amended in a pilot study among 543 Canadian adults (18–65 years old). The modified decisional balance scale consisted of 10 items (five for pros and five for cons) assessed using the 5-point response options ranging from 1 (not at all important) to 5 (very much) [14]. Despite widely used measurement, the validity of DB scale for Asian population is still limited. Cultural variation may influence the individual's exercise experience and “thus, when researchers and practitioners employ Western-based assessments with Asian populations by directly translating them without an appropriate validation, the process can be challenging” [15]. In this study, we translated the decisional balance scale into Malay language (DB-M) and confirmed its psychometric properties based on the modified version with ten items.

Given that, several statistics suggest that adult physical inactivity is a significant public health concern in Malaysia [16]. Apart from that, many studies have shown that the patterns of physical activity behaviours practised in college are likely to be sustained for a long time [17, 18]. Generally, all college students are adults with various responsibilities [19]. As such, they are more likely to maintain physical activity behaviour that they acquire during their college years and throughout their adulthood; therefore, the habit determines long-term health [17]. According to Sullum, Clark and King [20], health educators attempting to overcome the risk of exercise relapse should focus on reducing the cons and then shift focus to improve pros. Therefore, in this study, we aimed to examine the psychometric properties of the Malay version of the decisional balance scale (DB-M) and confirm its measurement invariance across gender among university students in Malaysia.

## Methods

### Participants

A total of 800 questionnaires were administered, and 750 (93.8% response rate) with complete responses to all the items were returned to the researchers. Hence, the final data set was 750 (female:  $n = 388$ , 51.7%; male:  $n = 362$ , 48.3%) with no missing data. The participants had a mean age of 20.2 years ( $SD = 1.2$ ) and they identified themselves as Malay (63.4%), Chinese (20.9%), Indian (9.7%), and others (6.0%), but they were all Malaysians with strong reading and speaking skills in the Malay language.

### Measures

The decisional balance scale was a self-report measure to evaluate the participants' perception of the pros (or perceived benefits) and cons (or perceived barriers) of maintaining a prevailing behaviour or embracing a new behaviour [14]. Examples of pros items included, “Physical activity would help me reduce tension or manage stress” while examples of cons items included “Physical activity would take too much of my time” [14]. The internal consistency of the modified version was satisfactory, with Cronbach's alpha of .79 for pros and .71 for cons [14].

### Questionnaire translation

The standard forward-backwards translation method was applied to translate the English version of the decisional balance scale into the Malay language [21]. First, a bilingual researcher

translated the DB scale from English to Malay to retain the content meaning and not word for word translation. Second, a local Malay who was bilingual in Malay and English back translated the Malay version to English. Third, these two translated versions were reviewed by a panel of five bilingual experts from the field of health psychology, sport sciences, physical education, and sports psychology. The items were matched with its corresponding items in the original English version and were further assessed to ascertain whether they were culturally appropriate for Malaysian populations. Finally, a sample of 30 undergraduate students was invited to assess the items' clarity, understanding, and comprehension. The students' feedback was similar to one another and required no revisions.

### Data collection

This study was approved by the Human Research Ethics Committee of the Universiti Sains Malaysia. The study was a cross-sectional design and implemented between September and December 2018. The purpose of the study and method of data collection were explained to the students at the end of their lecture sessions by the researchers. The students who consented to participate were given the information sheets containing the study's description before being asked to complete the questionnaire. Implied consent was considered to have been given when the participants voluntarily completed the DB-M questionnaire and returned it to the researchers. Time to complete the DB-M was at an average of 15 minutes. To report the test-retest reliability of the DB-M, 100 participants completed and returned the DB-M scale on day 14 after the first measurement.

### Statistical analysis

The Mardia multivariate normality of skew and kurtosis was initially conducted to examine the multivariate assumption of the data. The multivariate skew ( $p$ -value  $< .001$ ) and kurtosis ( $p$ -value  $< .001$ ) normality assumption was violated. Hence, the robust maximum likelihood estimator (MLR) was used because it provides robust estimates that are robust to non-normality violations [22, 23]. Many researchers have reported that MLR estimator can be used for ordinal CFA based models when the number of response levels for every item was equal to or greater than five [24, 25]. Also, a variable can be treated as a continuous measure when it is measured using a Likert scale with five or more response options [26]. Confirmatory factor analysis (CFA) along with measurement and structural invariances analyses were performed using Mplus version 8 to impose the hypothesised measurement structure on the data. The standardised factor loading of .40 or greater was used as a cut-off value to establish sufficient factor loading for all items of the DB-M scale. Items with factor loadings of less than .40 were considered to be uncertain items and were deleted with sufficient theoretical support [23].

Rather than applying the Chi-square goodness-of-fit ( $\chi^2$ ), Normed chi-square (NC) =  $\chi^2/df$  was originated because of the model chi-square sensitivity to the sample size [27]. Nevertheless, there was no precise guideline regarding the maximum acceptable values of NC. Following chi-square goodness-of-fit sensitivity to sample size and given that normed chi-square has no use in global fit testing, it is therefore not reported in this study. In the present study, the Standardised Root Mean Square Residual (SRMR), Root Mean Square Error of Approximation (RMSEA), Tucker-Lewis Index (TLI) and Comparative Fit Index (CFI) for estimating the validity of the CFA model were the only fit indices presented. This is in line with the recommendation of Brown [28], which stated that "the most universally accepted global goodness of fit indices are the SRMR, RMSEA, TLI, and CFI."

The recommended fit indices for a sample size of more than 250 with less than 12 used observe variables are: RMSEA of less than .07, SRMR of less than .08, and CFI or TLI of .95 or

higher [29]. The construct validity of the DB-M was further determined by computing the construct reliability (CR) and average variance extracted (AVE). CR was estimated using Raykov's method [30] in Mplus 8.0. The cut-off values were equal to or greater than .60 for CR [31] and .50 for AVE [32]. Discriminant validity was established by assessing the correlation between the factors, and a correlation coefficient of less than .85 suggests the adequate discriminant ability of the model [33]. Additionally, model re-specification was done by adding the residual covariances of items within the same factor. These error covariances reveal the assumption that the two corresponding items share some unexplained variances in common that is not specified in the model [27]. The model re-specification was based on the modification index (MI) values and carried out after the researchers obtained sufficient theoretical support.

The hierarchical test of measurement invariance was performed across gender using the recommended guidelines for establishing measurement invariance of the DB-M model [22, 26, 34]. A step by step restrictive constraints was imposed on the model parameters, and the differences in the model fit indices were examined. To examine measurement invariance, first, the configured invariance model needs to be established to compare it with the fit indices of other invariance models. In the configural invariance model, no equality restrictions were imposed on the model parameters across gender. Second, the weak invariance or metric invariance model was established and examined. In the weak invariance model, equality restrictions were imposed on the model factor loadings across gender to guarantee the similarity of the measurement scale across gender and make accurate comparisons. Third, the strong invariance model was established and examined. In the strong invariance model, equality restrictions were imposed on the factor loadings, as well as item intercepts across gender to guarantee the scale factors similarities across gender. Finally, a strict invariance model was established and examined. In the strict invariance model, equality restrictions were imposed on the factor loadings, item intercepts, and residual variances to determine that the items' variances of regression equations were invariant across gender.

The structural invariance of the model parameters was also investigated using the factor variance and factor covariance invariance, and the factor means invariance. The factor variance and factor covariance invariance were examined to ascertain the similarities of factors correlation across gender. In contrast, the factor means invariance was examined to ascertain the similarities of factors means differences across gender. In general, the structural invariance assessed the extent to which the male and female samples differ and not related to the measurement scale studied. For the present study, the recommended invariance cut-off values used were an absolute difference ( $\Delta$ ) of .01 or less for CFI ( $\Delta$ CFI) and TLI ( $\Delta$ TLI) and .015 for RMSEA ( $\Delta$ RMSEA) [26, 35, 36].

In this study, the internal consistency of the DB-M was evaluated based on Cronbach's alpha to compare it with the previous study's Cronbach's alpha [14, 37]. For test-retest reliability, the intraclass correlation coefficient (ICC) was computed using a subsample of the 100 participants' scores. Cronbach's alpha and ICC were computed using SPSS 24 and ICC values greater than .90 were considered to indicate excellent stability.

## Results

### Characteristics of the participants

The participants had a mean age of 20.2 (SD = 1.2); BMI of 21.6 (SD = 3.2); frequency of exercise/week of 3.0 (SD = 1.6); and duration of exercise/week of 63.2 minutes (SD = 37.7). The majority of the participants (98.7%) did not have any illnesses and were not on any medications.

## Measurement model of the DB-M

The hypothesised measurement model of the DB-M consisted of ten items and two factors (five items each). The initial model tested (Model 1) did not result in good fit indices (Table 1). However, all the standardised factor loadings were higher than .60 with  $p$ -value  $< .001$  (Fig 1). The initial model (Model 1) was improved by adding three pairs of correlated residuals within the same factor based on the modification indices. The results (Model 2) showed a good fit with the data (Table 1). Therefore, Model 2 was established with adequate fit indices and standardised factor loadings of .610–.799 (Fig 2).

## Construct reliability and discriminant validity

The CR values of the final DB-M model (model 2) were .850 for perceived benefits and .839 for perceived barriers. The AVE values for perceived benefits and perceived barriers were .542 and .538, respectively. For discriminant validity, the correlation coefficient between the two factors (.027) was less than the recommended value of .85, indicating sufficient discriminant validity. Table 2 presents the CR, AVE values, and the factors correlation coefficient of the final DB-M model.

## Measurement model of the DB-M for males and females

After establishing the overall baseline measurement model for all the data, the baseline measurement models for both males (Model 3) and females (Model 5) were specified. These two initial models did not result in a good fit with the data (Table 3). Therefore, both the male and female models have improved by adding correlated item residuals with the same factor for each model. The fit indices were adequate for both the final male model (Model 4: RMSEA = .052, CFI = .975, TLI = .964, SRMR = .040) and the final female model (RMSEA = .058, CFI = .965, TLI = .949, SRMR = .044). The final male model consisted of two added residual covariances (DB1 with DB2, and DB6 with DB7), and the final female model consisted of three added residual covariances (DB1 with DB2, DB9 with DB10, and DB3 with DB5). The baseline measurement models were not completely identical across groups (male and female). However, this differentiate in model specification should not affect model comparisons in other parameters in regard to invariance test [30, 31]. The standardised factor loadings of the final male and female models were .537–.840 (Fig 3) and .605–.858 (Fig 4), respectively.

## Measurement and structural invariance

The fit indices for both the male and female sample model were within the recommended cut-off values by Hair et al. [29]. Hence, these two models were established with all the items retained. Firstly, the male and female models were integrated to establish the configural model invariance with the same number of fixed and free factor loadings. The configural invariance model fit the data well across gender (Table 3). Secondly, the weak measurement invariance model was established with adequate fit indices (Table 3). The weak measurement invariance model, when compared with the non-restrictive model (the configural model), resulted in

**Table 1. Summary for DB-M model fit indices.**

Path model	RMSEA (90% CI)	CFI	TLI	SRMR
Model-1	.075 (.064, .086)	.939	.919	.044
Model-2 <sup>a</sup>	.047 (.034, .059)	.979	.969	.037

<sup>a</sup>Model-2 with correlated items residual; DB7 with DB6, DB2 with DB1, DB5 with DB3.

<https://doi.org/10.1371/journal.pone.0230644.t001>

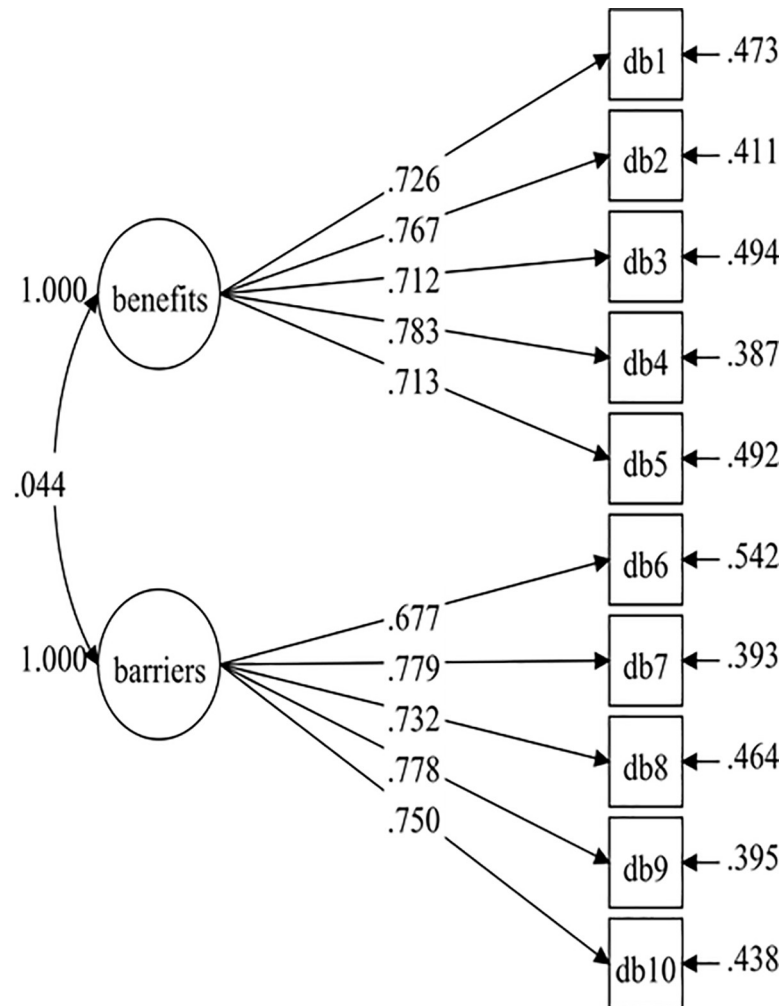


Fig 1. DB-M model (Model-1).

<https://doi.org/10.1371/journal.pone.0230644.g001>

adequate metric invariance differences across gender ( $\Delta\text{CFI} = -.005$ ,  $\Delta\text{TLI} = -.001$ ,  $\Delta\text{RMSEA} = 0$ ). This indicated that the items were interpreted similarly by both the male and female participants. Thirdly, the second most restrictive model (strong invariance model) was established. In the strong invariance model, equality restrictions were imposed on the factor loadings and item intercepts. These results ( $\Delta\text{CFI} = -.005$ ,  $\Delta\text{TLI} = -.001$ ,  $\Delta\text{RMSEA} = .001$ ) designated that the factor loadings and their intercepts remained invariant across gender. Finally, the highest most restrictive model (strict invariance model) was established. In the strict invariance model, equality restrictions were imposed on the factor loadings, item intercepts, and residual variances. These results ( $\Delta\text{CFI} = .014$ ,  $\Delta\text{TLI} = .010$ ,  $\Delta\text{RMSEA} = .006$ ) designated that the items' average scores remained invariant across gender.

The structural invariance of the DB-M was assessed using the factor variance and factor covariance invariance, and the factor means invariance. The factor variance and factor covariance invariance fit the data well ( $\text{CFI} = .957$ ,  $\text{TLI} = .953$ ,  $\text{SRMR} = .087$ ,  $\text{RMSEA} = .057$ ), and its differences with the less restrictive invariance model (strong invariance) are within the acceptable values ( $\Delta\text{CFI} = -.003$ ,  $\Delta\text{TLI} = -.002$ ,  $\Delta\text{RMSEA} = .001$ ). These results signified that the relationships among the two factors of the DB-M remained the same across gender. The factor

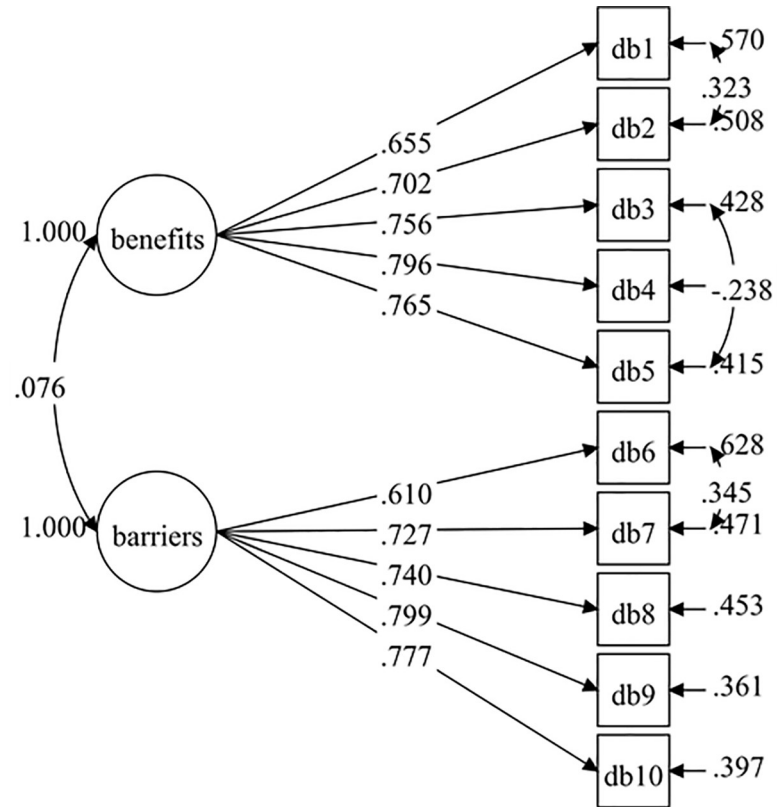


Fig 2. DB-M model (Model-2).

<https://doi.org/10.1371/journal.pone.0230644.g002>

means invariance fit indices were slightly below the recommended values (CFI = .939, TLI = .934, SRMR = .104, RMSEA = .067), and its differences with the less-restrictive model (factor variance and covariance) are slightly below the recommended values ( $\Delta$ CFI = -.018,  $\Delta$ TLI = -.019,  $\Delta$ RMSEA = .010). This result signals that the factor means are not invariant across gender.

### Internal consistency

The Cronbach’s alpha values were .857 for perceived benefits, and .859 for perceived barriers. The item-total correlation was .613–.716. According to Nunnally and Bernstein [38], the item-total correlation of greater than .30 signifies sufficient internal consistency, and each item contributes adequately to the measurement of its factor.

### Test-retest reliability

For test-retest reliability, 100 participants (female: 83, male: 17) volunteered to complete the DB-M again on day 14. The mean score for perceived benefits decreased from 21.0 (SD = 3.1)

Table 2. Composite reliability [10], average variance extraction (AVE), factor correlation and squared correlation for DB-M final model.

Variables	CR (95% CI)	AVE	1	2
1. Benefits	.850 (.824, .876)	.542	1	.027
2. Barriers	.839 (.817, .862)	.538		1

Correlation p-value = .120.

<https://doi.org/10.1371/journal.pone.0230644.t002>



**Table 3. The measurement model of DB-M baseline model fit results and tests of measurement and structure invariance.**

Models	CFI	TLI	RMSEA	SRMR	Model comparison	ΔCFI	ΔTLI	ΔRMSEA
Model-3 (male group: hypothesised)	.933	.912	.081	.047	-	-	-	-
Model-4 <sup>b</sup> (male group: re-specified)	.975	.964	.052	.040	-	-	-	-
Model-5 (female group: hypothesised)	.927	.904	.079	.054	-	-	-	-
Model-6 <sup>c</sup> (female group: re-specified)	.965	.949	.058	.047	-	-	-	-
Model-7 (configural)	.970	.957	.055	.044	-	-	-	-
<b>Measurement invariance</b>								
Model-8 (weak)	.965	.956	.055	.057	8 versus 7	-.005	-.001	0
Model-9 (strong)	.960	.955	.056	.062	9 versus 8	-.005	-.001	.001
Model-10 (strict)	.946	.945	.062	.065	10 versus 9	.014	-.010	.006
<b>Structural invariance</b>								
Model-11 (factor variance and covariance)	.957	.953	.057	.087	11 versus 9	-.003	-.002	.001
Model-12 (factor variance, covariance and factor mean)	.939	.934	.067	.104	12 versus 11	-.018	-.019	.010

<sup>b</sup>Adding residual covariances between item DB1 with DB2, DB7 with DB6.

<sup>c</sup>Adding residual covariances between DB1 with DB2, DB10 with DB9, DB5 with DB3. Model-1 and Model-2 are reported in Table 1.

<https://doi.org/10.1371/journal.pone.0230644.t003>

to 19.8 (SD = 3.3), with an ICC value of .979 (95% CI, .969, .986, p-value < .001). The mean score for perceived barriers increased from 10.4 (SD = 3.2) to 11.6 (SD = 3.9), with an ICC value of .960 (95% CI, .941, .973, p-value < .001). Based on gender difference: the ICC for perceived benefits was .975 (95% CI: .933, .991, p-value < .001) for males and .980 (95% CI: .969, .987, p-value < .001) for females, and the ICC for perceived barriers was .944 (95% CI: .854, .980, p-value < .001) for males and .962 (95% CI: .941, .975, p-value < .001) for females. These ICC values revealed that the DB-M had an excellent stability over time [39].

## Discussion

The present study aimed to translate the English version of the DB scale into Malay (DB-M) and then evaluate its validity with university students using CFA. This study presented evidence for adequate psychometric properties of the DB-M scale consistent with previous studies [14, 37]. The DB-M scale fit the data well, and the final model retained all ten items with strong factor loadings (above .60) on their respective factors. Furthermore, the results provided substantial evidence for measurement and structural invariance across gender.

The DB-M demonstrated adequate construct validity and discriminant validity. The CR and variance extracted of the DB-M, or the degree to which the items revealed their respective factors, exceeded the prescribed values of .60 for CR [31] and .50 for AVE [32]. These results showed that the DB-M had adequate construct validity, and all the items accurately estimate their respective factors. For discriminant validity of the DB scale, the correlation coefficient between the two factors (.027) was less than the prescribed value of .85 [33]. These results showed sufficient discriminant validity of the DB scale, with each factor explaining distinct information from the other factor.

According to the TTM, individuals need many struggles at the behavioural change to be successful [40]. The decisional balance is among the psychological constructs of the TTM that influences exercise behaviour [41]. Therefore, it is necessary to ascertain an instrument that reliably assesses the decisional balance constructs. The DB-M model tested in the present study reveals sufficient internal consistency. The Cronbach's alpha values of .857 for perceived benefits and .859 for perceived barriers matches the previous studies Cronbach's of .79 and .71 [14], and .87 and .70 [37], for perceived benefits and perceived barriers, respectively.

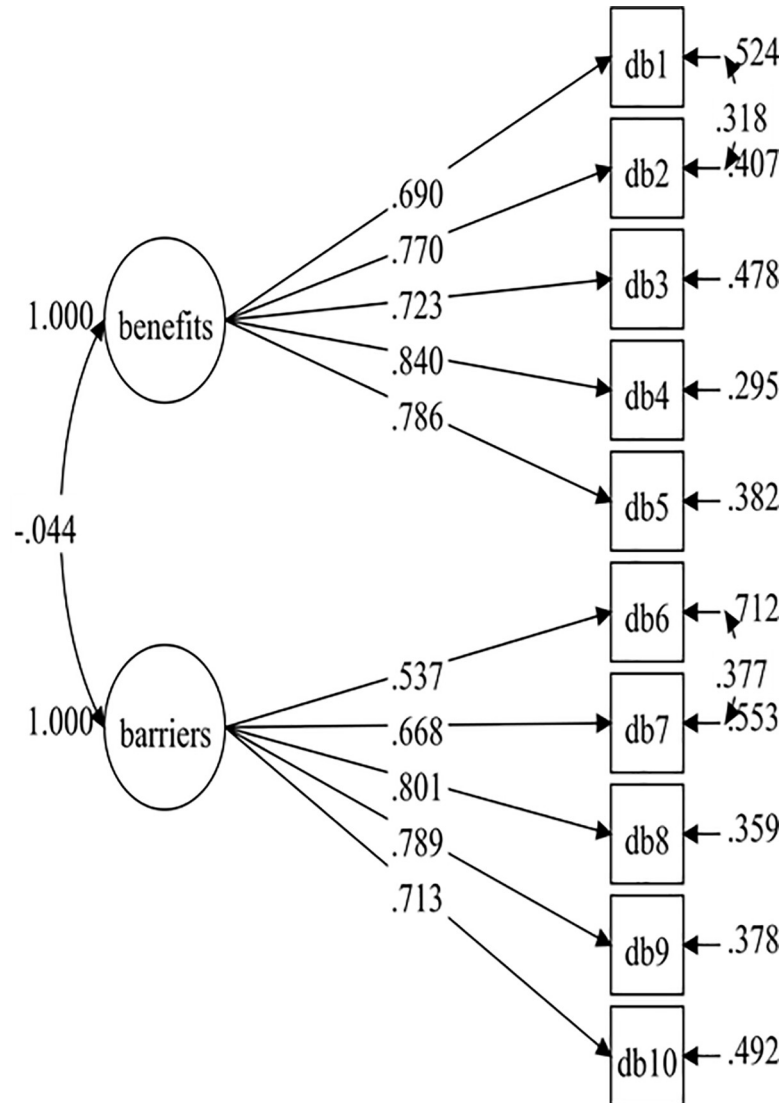
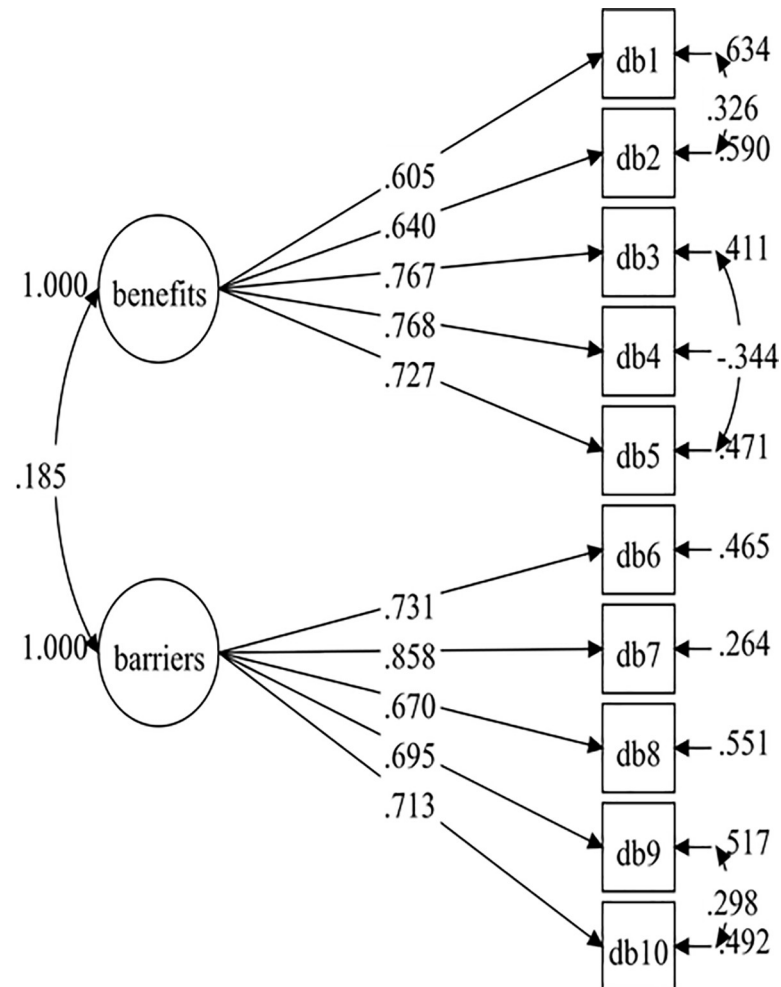


Fig 3. Male DB-M model (Model-4).

<https://doi.org/10.1371/journal.pone.0230644.g003>

Additionally, the ICC values for test-retest reliability (.979 for perceived benefits and .960 for perceived barriers) computed in this study indicate that the DB-M scale has excellent stability over time [39]. A previous study reported a 2-weeks test-retest reliability coefficient of .91 and .89 for perceived benefits and perceived barriers, respectively [41].

The present study investigated the measurement invariance of the DB-M across gender and ascertained that all the invariance requirements were met [26, 42]. These results signified that the male and female samples had similar understandings of all the ten items in the DB-M, which is vital when making valid comparisons between the male and female students' decisional balance to exercise. Furthermore, the structural invariance of the DB-M was desirable for factor variance and covariance, but the factor means were not desirable (variant) across gender. These results indicate that the relationships between the factors remain the same across gender, but the means of the factors varies across the gender. A previous study by McAuley et al. [43] reported that males have higher confidence than females in completing exercise prescription.



**Fig 4. Female DB-M model (Model-6).**

<https://doi.org/10.1371/journal.pone.0230644.g004>

Residual covariances may be specified as a means to examine hypotheses regarding shared sources of variability over and beyond the factors [31]. In the present study, covariances between residuals' items within the same factor were added to improve the model fit indices. These modifications on the hypothesized DB-M model were determined based on the MI values reported in Mplus output and after sufficient theoretical support was established by the researchers. Covariance between residuals for items DB1 (Physical activity would help me reduce tension or manage stress) and DB2 (I would feel more confident about my health by getting physical activity) were added to both the final overall model, male model, and female model. This is reasonable, as previous studies reported a strong relationship between mental health, stress, mood and physical exercise [44]. Other covariances added were the error residual for DB6 (I am too tired to get physical activity because of my other daily responsibilities) and DB7 (Physical activity would take too much of my time) for the final overall model and male model, the error residual for DB3 (I would sleep better) and DB5 (Physical activity would help me control my weight) for the final overall model and female model, and the error residual for DB9 (I'd worry about looking awkward if others saw me being physically active) and DB10 (Getting physical activity would cost too much money) for the final female model only. A better fitting model was established after adding these residual covariances. In social psychological

research, when these covariances make a substantive sense, they should be included in the model [45, 46].

In this study, some study limitations exist, and recommendations for future research are needed. First, although the sample size for this study is considered to be large, caution must be given to the generalisability of the results, as the data were collected from a single university. Second, the self-reported measure was used to assess the student's decisional balance to exercise. Self-reported measures have been associated with response bias, which could reduce the accuracy of the data obtained. Nevertheless, the questions were answered anonymously, and the participants were assured of their confidentiality. Also, future research should examine the replicability of the DB-M with a more diverse Malaysians population of different ages, education levels, occupations, and health conditions.

## Conclusion

In the present study, the DB-M was shown to have adequate psychometric properties and can be administered to evaluate decisional balance for exercise behaviour among university students. All the items were retained with strong factor loadings. Also, the scale displayed sufficient measurement invariance (configural, weak, strong, and strict) and structural invariance (factor variance and covariance). These findings illustrate that DB-M can be used to make valid comparisons across gender.

## Supporting information

### S1 Data.

(PDF)

### S2 Data.

(SAV)

## Author Contributions

**Conceptualization:** Garry Kuan, Yee Cheng Kueh.

**Data curation:** Garry Kuan, Yee Cheng Kueh.

**Formal analysis:** Abdulwali Sabo, Yee Cheng Kueh.

**Funding acquisition:** Garry Kuan, Yee Cheng Kueh.

**Investigation:** Garry Kuan, Yee Cheng Kueh.

**Methodology:** Garry Kuan, Yee Cheng Kueh.

**Project administration:** Garry Kuan.

**Resources:** Garry Kuan.

**Software:** Yee Cheng Kueh.

**Supervision:** Garry Kuan, Yee Cheng Kueh.

**Validation:** Garry Kuan, Yee Cheng Kueh.

**Writing – original draft:** Abdulwali Sabo.

**Writing – review & editing:** Garry Kuan, Abdulwali Sabo, Sukanlaya Sawang, Yee Cheng Kueh.

## References

1. Sabo A, Kueh YC, Kuan G. Psychometric properties of the Malay version of the self-efficacy for exercise scale. *PLoS ONE*. 2019; 14(5):e0215698. <https://doi.org/10.1371/journal.pone.0215698> PMID: 31050664
2. Rhodes RE, Plotnikoff RC, Courneya KS. Predicting the physical activity intention–behavior profiles of adopters and maintainers using three social cognition models. *Ann Behav Med*. 2008; 36(3):244–52. <https://doi.org/10.1007/s12160-008-9071-6> PMID: 19057976
3. Booth FW, Laye MJ, Lees SJ, Rector RS, Thyfault JP. Reduced physical activity and risk of chronic disease: the biology behind the consequences. *Eur J Appl Physiol*. 2008; 102(4):381–90. <https://doi.org/10.1007/s00421-007-0606-5> PMID: 17987311
4. WHO. Global health risks: mortality and burden of disease attributable to selected major risks: Geneva: World Health Organization; 2009.
5. Poh BK, Safiah M, Tahir A, Siti Haslinda N, Siti Norazlin N, Norimah A, et al. Physical activity pattern and energy expenditure of Malaysian adults: findings from the Malaysian adult nutrition survey (MANS). *Malays J Nutr*. 2010; 16(1):13–37. PMID: 22691851
6. Spencer L, Adams TB, Malone S, Roy L, Yost E. Applying the transtheoretical model to exercise: a systematic and comprehensive review of the literature. *Health Promot Pract*. 2006; 7(4):428–443. <https://doi.org/10.1177/1524839905278900> PMID: 16840769
7. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. *J Consult Clin Psychol*. 1983; 51(3):390 <https://doi.org/10.1037//0022-006x.51.3.390> PMID: 6863699
8. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot*. 1997; 12(1):38–48. <https://doi.org/10.4278/0890-1171-12.1.38> PMID: 10170434
9. Gorczynski P, Faulkner G, Greening S, Cohn T. Exploring the construct validity of the transtheoretical model to structure physical activity interventions for individuals with serious mental illness. *Psychiatr Rehabil J*. 2010; 34(1):61–64. <https://doi.org/10.2975/34.1.2010.61.64> PMID: 20615847
10. Tuah NA, Amiel C, Qureshi S, Car J, Kaur B, Majeed A. Transtheoretical model for dietary and physical exercise modification in weight loss management for overweight and obese adults. *Cochrane Database Syst Rev*. 2011;(10). <https://doi.org/10.1002/14651858.CD008066.pub2> PMID: 21975777
11. Janis IL, Mann L. Decision making: a psychological analysis of conflict, choice, and commitment. Free press; 1977.
12. Nigg CR, Courneya KS, Estabrooks PA. Maintaining attendance at a fitness center: an application of the decision balance sheet. *Behav Med*. 1997; 23(3):130–137. <https://doi.org/10.1080/08964289709596369> PMID: 9397285
13. Marcus BH, Rakowski W, Rossi JS. Assessing motivational readiness and decision making for exercise. *Health Psychol*. 1992; 11(4):257–261. <https://doi.org/10.1037//0278-6133.11.4.257> PMID: 1396494
14. Plotnikoff RC, Blanchard C, Hotz SB, Rhodes R. Validation of the decisional balance scales in the exercise domain from the transtheoretical model: A longitudinal test. *Meas Phys Educ Exerc Sci*. 2001; 5(4):191–206. [https://doi.org/10.1207/S15327841MPEE0504\\_01](https://doi.org/10.1207/S15327841MPEE0504_01)
15. Oei TPS, Sawang S, Goh YW, Mukhtar F. Using the depression anxiety stress scale 21 (DASS-21) across cultures. *Int J Psychol*. 2013; 48(6), 1018–1029. <https://doi.org/10.1080/00207594.2012.755535> PMID: 23425257
16. Ying C, Kuay LK, Huey TC, Hock LK, Hamid HAA, Omar MA, et al. Prevalence and factors associated with physical inactivity among Malaysian adults. *Southeast Asian J Trop Med Public Health*. 2014; 45(2):467–480. PMID: 24968689
17. Sparling PB, Snow TK. Physical activity patterns in recent college alumni. *Res Q Exerc Sport*. 2002; 73(2):200–205. <https://doi.org/10.1080/02701367.2002.10609009> PMID: 12092895
18. Fish C, Nies MA. Health promotion needs of students in a college environment. *Public Health Nurs*. 1996; 13(2):104–111. <https://doi.org/10.1111/j.1525-1446.1996.tb00227.x> PMID: 8936243
19. Calfas K, Sallis J, Lovato C, Campbell J. Physical activity and its determinants before and after college graduation. *Med Exerc Nutr Health*. 1994; 3(323–334):28–37.
20. Sullum J, Clark MM, King TK. Predictors of exercise relapse in a college population. *J Am College Health*. 2000; 48(4):175–180. <https://doi.org/10.1080/07448480009595693> PMID: 10650735
21. Brislin RW. Back-translation for cross-cultural research. *J Cross Cult Psychol*. 1970; 1(3):185–216. <https://doi.org/10.1177/135910457000100301>
22. Muthén L, Muthén B. Mplus user's guide (8<sup>th</sup> ed.). Los Angeles: Muthén and Muthén.; 1998.

23. DeVon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, et al. A psychometric toolbox for testing validity and reliability. *J Nurs Scholarsh*. 2007; 39(2):155–164. <https://doi.org/10.1111/j.1547-5069.2007.00161.x> PMID: 17535316
24. Raykov T. (2012). Scale construction and development using structural equation modeling. In Hoyle R. H. (Ed.), *Handbook of structural equation modeling* (pp. 472–492). New York, NY: Guilford Press.
25. Rigdon E. E. (1998). Structural equation modeling. In Marcoulides G. A. (Ed.), *Modern methods for business research* (pp. 251–294). Mahwah, NJ: Erlbaum.
26. Wang J, Wang X. *Structural equation modeling: Applications using Mplus*. United States: John Wiley & Sons.;2012.
27. Kline R. *Principles and Practice of Structural Equation Modeling*, 3rd edn. New York: Guilford Press;2011.
28. Brown T. A. *Confirmatory factor analysis for applied research*: Guilford Publications;2015
29. Hair JF, Black WC, Babin B, Anderson R, Tatham R. *Multivariate data analysis* ( 7th ed.). New Jersey: Pearson Prentice Hall;2010.
30. Raykov T, Marcoulides GA. Scale reliability evaluation under multiple assumption violations. *Struct Equ Modeling*. 2016; 23(2):302–313. <https://doi.org/10.1080/10705511.2014.938597>
31. Tseng W-T, Dörnyei Z, Schmitt N. A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Appl Linguist*. 2006; 27(1):78–102. <https://doi.org/10.1093/applin/ami046>
32. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res*. 1981; 18(1):39–50. <https://doi.org/10.1177/002224378101800104>
33. Browne M, & Cudeck R. *Confirmatory factor analysis for applied research: alternative ways of assessing model fit*. New York NY: Guilford; 2006.
34. Byrne BM. *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York: Routledge; 2013.
35. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling*. 2002; 9(2):233–255. [https://doi.org/10.1207/S15328007SEM0902\\_5](https://doi.org/10.1207/S15328007SEM0902_5)
36. Kimber M, Rehm J, Ferro MA. Measurement invariance of the WHODAS 2.0 in a population-based sample of youth. *PLoS ONE*. 2015; 10(11):e0142385. <https://doi.org/10.1371/journal.pone.0142385> PMID: 26565410
37. Hausenblas HA, Nigg CR, Downs DS, Fleming DS, Connaughton DP. Perceptions of exercise stages, barrier self-efficacy, and decisional balance for middle-level school students. *J Early Adolesc*. 2002; 22(4):436–454. <https://doi.org/10.1177/027243102237191>
38. Nunnally J, Bernstein I. *Psychometric theory*. ( 3rd ed.) McGrawHill: New York, NY, USA. 1994:
39. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016; 15(2):155–163. <https://doi.org/10.1016/j.jcm.2016.02.012> PMID: 27330520
40. Marcus BH, Selby VC, Niaura RS, Rossi JS. Self-efficacy and the stages of exercise behavior change. *Res Q Exerc Sport*. 1992; 63(1):60–66. <https://doi.org/10.1080/02701367.1992.10607557> PMID: 1574662
41. Kim YH. Application of the transtheoretical model to identify psychological constructs influencing exercise behavior: a questionnaire survey. *Int J Nurs Stud*. 2007; 44(6):936–944. <https://doi.org/10.1016/j.ijnurstu.2006.03.008> PMID: 16698024
42. Meredith W, Teresi JA. An essay on measurement and factorial invariance. *Med Care*. 2006: S69–S77. PMID: 17060838
43. McAuley E, Courneya KS, Lettunich J. Effects of acute and long-term exercise on self-efficacy responses in sedentary, middle-aged males and females. *Gerontologist*. 1991; 31(4):534–542. <https://doi.org/10.1093/geront/31.4.534> PMID: 1894158
44. Mehrad A. Relationship Between Stress, Physical Exercise, and Mental Health Among Athletes. *Journal of Educational, Health and Community Psychology*. 2017; 6(1): 18–23.
45. Sorbom D, Joreskog K. *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific Software. 1993:
46. Enders CK, Tofighi D. Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological methods*. 2007; 12(2): 121. PMID: 17563168