

Exploring the intention to walk: a study on undergraduate students using item response theory and theory of planned behaviour

Carla Galluccio, Rosa Fabbricatore, Daniela Caso

1. Introduction

Physical activity is one of the most basic human functions, and it is an important foundation of health throughout life. Physical activity apports benefit on both physical and mental health, reducing the risk of several diseases and lowering stress reactions, anxiety and depression (Penedo, Dahn, 2005). More specifically, physical activity is defined as “any bodily movement produced by skeletal muscles that require energy expenditure” (WHO, 2018), including in this definition several activities. Among them, walking has been shown to improve physical and mental well-being in every age group.

In this regard, the World Health Organization has suggested taking a goal of about 10,000 steps per day. However, achieving this goal may be difficult for many. For this reason, Tudor-Locke and Bassett (2004) proposed to lower the threshold at least 7,000 steps a day. Despite that, insufficient walking among university students has been increasingly reported (Sun et al., 2015), requiring walking promotion intervention (e.g. Caso et al., 2020). In order to do this, dividing students based on their intention to walk might be useful, since intention is considered the best predictor of behaviour. In this regard, the main theoretical framework used to explain physical activity is the Theory of Planned Behaviour (TPB; Ajzen, 1991).

In this theory, behavioural intention is determined by three factors. The first predictor of intention is the attitude toward behaviour (both affective and instrumental; see Lowe, Eves, Carroll, 2002 for details), that is the evaluation of the behaviour as favourable or unfavourable. The second factor are subjective norms, which refer to individual’s beliefs about whether an important person or group of people approved or not the behaviour. Finally, the third antecedent of intention is the perceived behavioural control (PBC), which can be defined as the individual’s perception of the easiness or difficulty of performing the behaviour (Ajzen, 1991).

Herein, we decided to extend the traditional TPB model adding two additional variables as walking intention’s predictors, namely self-identity and risk perception. The former is defined as salient and prominent aspects of one’s self-perception, whereas the latter refers to the subjective judgement about the severity of a risk. In this regard, some studies have shown that self-identity emerged as a significant predictor of intention to walk in different population (e.g. Ries et al., 2012). Besides, past research (e.g. Stephan et al., 2011) has also shown that risk perception could affect physical activity motivation and behaviour. For these reasons, may be reasonable to suppose that these predictors could be significant also for university students.

In this work, we investigated the university students’ intention to walk by exploiting Item Response Theory (IRT) models (Bartolucci, Bacci, Gnaldi, 2015). In particular, we inspected the predictors of intention by means of Rating Scale Graded Response Model (RS-GRM; Muraki, 1990). Afterwards, we used the Latent Class RS-GRM (Bacci, Bartolucci, Gnaldi, 2014) to divide students according to their intention to walk, including predictors’ scores as covariates.

Carla Galluccio, University of Florence, Italy, carla.galluccio@unifi.it, 0000-0003-1154-3601

Rosa Fabbricatore, University of Naples Federico II, Italy, rosa.fabbricatore@unina.it, 0000-0002-4056-4375

Daniela Caso, University of Naples Federico II, Italy, daniela.caso@unina.it, 0000-0002-6579-963X

FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Carla Galluccio, Rosa Fabbricatore, Daniela Caso, *Exploring the intention to walk: a study on undergraduate students using item response theory and theory of planned behaviour*, pp. 153-158, © 2021 Author(s), CC BY 4.0 International, DOI 10.36253/978-88-5518-304-8.30, in Bruno Bertaccini, Luigi Fabbris, Alessandra Petrucci, *ASA 2021 Statistics and Information Systems for Policy Evaluation. Book of short papers of the opening conference*, © 2021 Author(s), content CC BY 4.0 International, metadata CCO 1.0 Universal, published by Firenze University Press (www.fupress.com), ISSN 2704-5846 (online), ISBN 978-88-5518-304-8 (PDF), DOI 10.36253/978-88-5518-304-8

2. Participants and procedure

Data was collected administrating an online self-report questionnaire to undergraduate students enrolled in the Psychology course at Federico II University of Naples. The final sample included $N = 146$ students.

Regarding the questionnaire, for the traditional TPB variables we adapted the scale proposed by Ajzen (2002): *intention* was assessed by 3 items (e.g. “I intend to walk 7,000 steps a day”); *subjective norms* were assessed by 5 items (e.g. “Most people who are important to me think that I should do 7,000 steps a day”); *PBC* was assessed by 4 items (e.g. “Doing 7,000 steps a day is under my control”). For these variables we used a 7-point Likert response scale (1 = strongly disagree to 7 = strongly agree). About *attitude*, it was assessed by 8 items on a semantic differential scale, with 4 items for both instrumental and affective attitude (e.g. “disadvantageous-advantageous” and “unpleasant-pleasant”, respectively). On the other hand, we assessed *self-identity* using 4 items (1 = strongly disagree to 7 = strongly agree response scale), e.g. “I think of myself as a physically active subject” (Fishbein, Ajzen, 2010). Finally, *risk perception* was assessed by 6 items (1 = not at all to 7 = very much response scale), e.g. “I think I am personally exposed to the risk of heart disease” (Petrillo, Caso, Donizzetti, 2004).

3. Statistical analysis

IRT model for ordinal polytomous items was carried out for measuring all the TPB variables. In particular, the analysis made up of two steps. Firstly, we estimated the predictors of intention exploiting the RS-GRM as the best model selected among several others according to the BIC index (Schwarz, 1978). For the attitude variable we carried out a bi-dimensional RS-GRM since attitude consists of both instrumental and affective dimensions, whereas for the other variables we used a uni-dimensional RS-GRM. In the second step of our analysis we divided students according to their intention to walk by using a Latent Class RS-GRM. We considered the TPB predictors of intention to walk as individual covariates by using the scores obtained in the first step of the analysis. The analyses were computed using R statistical software.

Let Y_{ij} the response of individual i (with latent trait θ_i) to a polythomous item j with l_j response categories indexed from 0 to $l_j - 1$, the formulation of the GRM (Samejima, 2016) can be expressed as:

$$g_x[P(Y_{ij} = x|\theta_i)] = \log \frac{P(Y_{ij} \geq x|\theta_i)}{P(Y_{ij} < x|\theta_i)} = \gamma_j(\theta_i - \beta_{jx}), \quad j = 1, \dots, r, \quad x = 1, \dots, l_j - 1, \quad (1)$$

where $g_x(\cdot)$ is the global logit link function. The item parameters γ_j and β_{jx} represent the *discrimination* and the item-step *difficulty* parameter, respectively. It is worth noting that in this context a useful tool to evaluate the goodness of an item or a test as a whole is the Fisher information (Bartolucci, Bacci, Gnaldi, 2015).

A multidimensional extension of IRT models has been proposed to taking into account the correlation between multiple latent traits (Reckase, 2009). Therefore, each subject i is described by a vector of latent variables $\theta_i = (\theta_{i1}, \dots, \theta_{iD})$, where D indicates the number of dimensions in the model. According to the between-item multidimensional approach, each item measures only one latent trait. In particular, for the GRM we have:

$$\log \frac{P(Y_{ij} \geq x|\theta_i)}{P(Y_{ij} < x|\theta_i)} = \gamma_j \left(\sum_{d=1}^D \delta_{jd} \theta_{id} - \beta_{jx} \right) \quad (2)$$

where δ_{jd} is a dummy variable indicating if the item j measures the latent trait d ($\delta_{jd} = 1$) or not ($\delta_{jd} = 0$), with $d = 1, \dots, D$.

In this vein, the RS-GRM, adopted in the first step of our analysis, represents a constrained version of the GRM in which β_{jx} is expressed in an additive way, namely $\beta_{jx} = \beta_j + \tau_x$. According to this formulation, items may have different general difficulty level (β_j), but equal response category difficulty level (τ_x).

In the second step of the analysis we exploited a Latent Class IRT model, a semi-parametric extension of the IRT model allows to detecting sub-populations of individuals that are homogeneous with respect to the latent trait. The latter is represented through a discrete distribution with ξ_1, \dots, ξ_k support points defining k latent classes with weights π_1, \dots, π_k . It is worth noting that $\pi_c = P(\Theta = \xi_c)$ represents the prior probability of belonging to the latent class c ($c = 1, \dots, k$) with $\sum_{c=1}^k \pi_c = 1$ and $\pi_c \geq 0$. The discreteness of the latent trait leads to express the manifest distribution of the response vector $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{ir})'$ as:

$$P(\mathbf{Y}_i) = \sum_{c=1}^k P(\mathbf{Y}_i|\xi_c)\pi_c \quad (3)$$

where $P(\mathbf{Y}_i|\xi_c) = \prod_{j=1}^r P(Y_{ij} = x|\xi_c)$ due to the *local independence* assumption.

In particular, in this work we refers to the RS-GRM parameterisation, selected again as the best model by the BIC, so that:

$$\ln[P(Y_{ij} = x|\xi_c)] = \ln \frac{P(Y_{ij} \geq x|\xi_c)}{P(Y_{ij} < x|\xi_c)} = \gamma_j[\theta_i - (\beta_j + \tau_x)]. \quad (4)$$

When a vector of individual covariates \mathbf{Z}_i is considered, as in our analysis, the weight π_c is replaced with the individual weight $\pi_{ci} = P(\Theta = \xi_c|\mathbf{Z}_i = \mathbf{z}_i)$. About that, according to the global logit formulation, possible only when latent classes are ordered with respect to the latent trait, we have:

$$\ln \frac{\pi_{ci} + \pi_{(c+1)i} + \dots + \pi_{ki}}{\pi_{1i} + \pi_{2i} + \dots + \pi_{(c-1)i}} = \beta_{0c} + \mathbf{z}'_i \beta_1, \quad (5)$$

where β_{0c} is the class-specific constant term and β_1 is the vector of regression coefficients describing the effect of individual covariates (Dayton, Macready, 1988).

The estimation of the model parameters is obtained using the *Maximum Marginal Likelihood* (MML) approach (see Bartolucci, Bacci, Gnaldi, 2015 for details). The number of latent classes k was chosen by comparing the fit of models using different values of k .

4. Results

The latent trait analysis in the first step pointed to a good test Fisher information for all the predictors of the intention to walk we considered (see Figures 1 and 2). In particular, items measuring PBC, self-identity and attitude are maximally informative for students with low levels of the latent trait; whereas the test information curve for the risk perception is shifted on the right (greater information for high levels of the latent trait).

Regarding the Latent Class IRT model, the BIC indicated the RS-GRM with $k = 4$ number of classes as the best model. The standardised support points and the average of the individual weights π_{ci} are reported in Table 1. Looking at support points, we notice that latent classes are increasing ordered according to the levels of intention to walk 7,000 steps a day. On the other hand, the average weights indicated that Class 3 is the largest one, followed by Class 2. Thus, the majority of the students reported a medium level of intention to walk 7,000 steps a day.

Besides, in Table 2 we reported the TPB predictors that significantly affect the class weights. To estimate this effect, we adopted the global logit specification (see Equation 5) since the

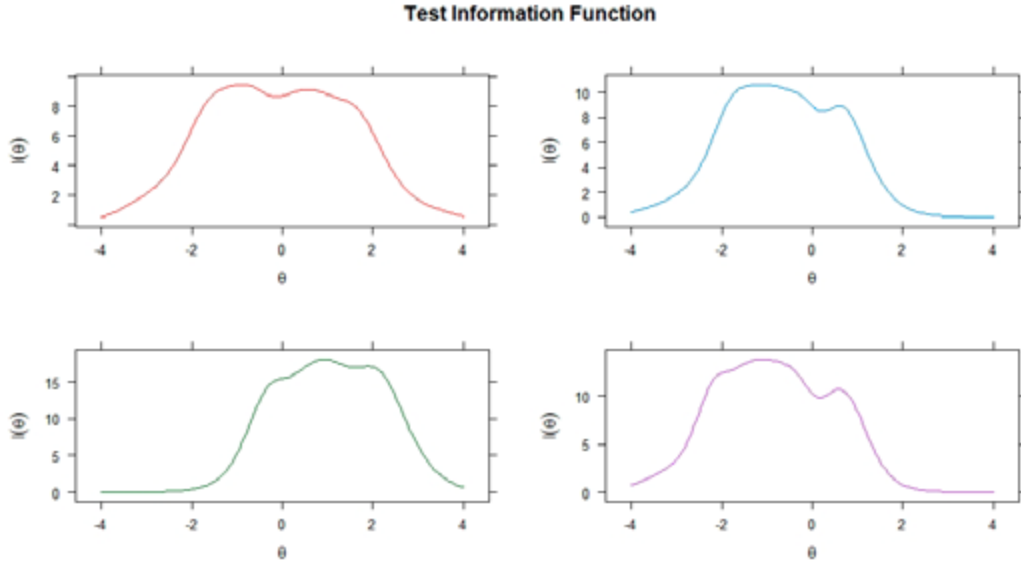


Figure 1: Test Fisher information curve for the predictors of the intention to walk: subjective norms (red line), PBC (blue line), risk perception (green line), and self-identity (purple line).

Table 1: Standardised support points and average class weights for the Latent Class RS-GRM with covariates.

	Latent class			
	1	2	3	4
Intention to walk 7,000 steps a day	-1.79	-0.62	0.36	1.65
<i>Average weights</i>	<i>0.13</i>	<i>0.26</i>	<i>0.46</i>	<i>0.15</i>

Table 2: Regression coefficient ($\hat{\beta}_1$), standard error (\hat{se}), t-value, and p-value for the individual covariates.

Covariate	$\hat{\beta}_1$	\hat{se}	t-value	p-value
Affective Attitude	0.97	0.21	4.62	0.00
Subjective Norms	0.38	0.19	2.04	0.04
Self-identity	0.42	0.18	2.31	0.02

support points were increasingly ordered. We removed from the final model all the variables resulted not significant for $\alpha = 0.10$, namely instrumental attitude, PBC, and risk perception.

We can conclude that the most significant covariate affecting positively the student's intention to walk 7,000 steps a day is affective attitude ($\hat{\beta}_1 = 0.97$, p-value < 0.01), followed by self-identity ($\hat{\beta}_1 = 0.42$, p-value < 0.05) and subjective norms ($\hat{\beta}_1 = 0.38$, p-value < 0.05).

5. Discussion and conclusion

The present study aimed to detect homogeneous groups of university students according to their intention to walk exploiting IRT models. We found that students could be divided into four ordered classes: Class 1 is made up of students with the lowest intention to walk, whereas Class 4 includes students with the highest intention to walk 7,000 steps a day. Besides,

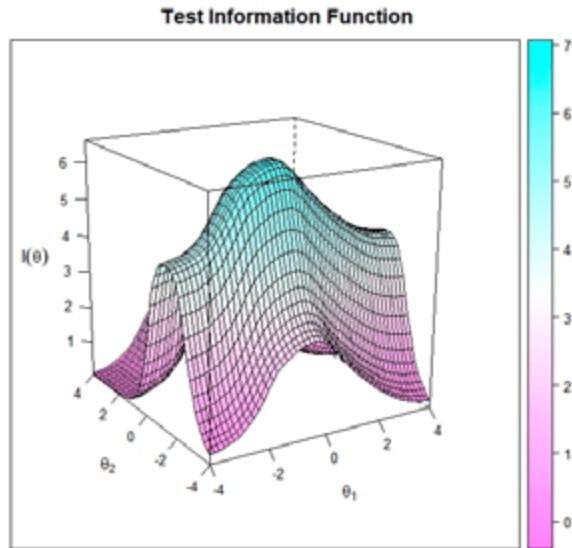


Figure 2: Test Fisher information curve for the attitude variable: θ_1 refers to the affective dimension, whereas θ_2 to the instrumental one.

results showed that the best predictors of intention to walk were affective attitude, subjective norms and self-identity. In contrast, instrumental attitude, risk perception and PBC were not significant. Regarding affective and instrumental attitudes, several studies on health behaviours have shown that affective attitude was a strong predictor of intention, often at the expense of instrumental attitude (e.g. Lowe, Eves, Carroll, 2002). Usually, health promotion programmes emphasised the instrumental benefits of physical activity, such as improved health, which are not immediately apparent to the individual due to the delay between doing physical activity and its results. Conversely, affective components of physical activity, such as its pleasant nature, are immediate consequences of involvement. Concerning subjective norms, results showed a moderate and positive influence on students' intention to walk. This finding is consistent with those in literature (Wing Kwan, Bray, Martin Ginis, 2009), where it is supposed that social influences on physical activity intention were stronger among younger populations. Besides, as we expected, self-identity resulted as a significant predictor of intention to walk in university students. In fact, according to the literature (e.g. Ries et al., 2012), our results are consistent with an interpretation that who identify themselves as physically active persons are more likely to practise regular physical activity. Finally, regarding risk perception and PBC, we found that in our model these variables are not significant. About risk perception, it is reasonable to suppose that the perception of the riskiness correlated with the physical inactivity, such as physical and mental diseases, is more likely in older than younger populations. On the contrary, the finding that PBC is not a significant predictor of intention was quite surprising. We speculate that university commitment leads students to not fully perceived the extent of their control on other activities, such as physical activity, especially during the first year.

In conclusion, we believe that the Latent Class IRT models represent a useful statistical tool for dividing students according to their intention to walk in order to define a more tailored walking promotion programmes. Indeed, we could support students in Class 4 in maintaining their intention to walk, whereas a different walking promotion intervention could be implemented for students in Class 1, focusing on the TPB variables that resulted as significant predictors.

References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, **50**(2), pp. 179–211.
- Ajzen, I. (2002). Constructing a TPB questionnaire: Conceptual and methodological considerations.
- Bacci, S., Bartolucci, F., Gnaldi, M. (2014). A class of Multidimensional Latent Class IRT models for ordinal polytomous item responses. *Communications in Statistics-Theory and Methods*, **43**(4), pp. 787–800.
- Bartolucci, F., Bacci, S., Gnaldi, M. (2015). *Statistical analysis of questionnaires: A unified approach based on R and Stata*. CRC Press, Boca Raton, (FL).
- Caso, D., Carfora, V., Capasso, M., Oliano, D., Conner, M. (2020). Using Messages Targeting Psychological versus Physical Health Benefits to Promote Walking Behaviour: A Randomised Controlled Trial. *Applied Psychology: Health and Well-Being*.
- Dayton, C.M., Macready, G.B. (1988). Concomitant-variable latent-class models. *Journal of the American statistical association*, **83**(401), pp. 173–178.
- Fishbein, M., Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. Taylor & Francis, New York, (NY).
- Lowe, R., Eves, F., Carroll, D. (2002). The influence of affective and instrumental beliefs on exercise intentions and behavior: A longitudinal analysis. *Journal of Applied Social Psychology*, **32**(6), pp. 1241–1252.
- Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement*, **14**(1), pp. 59–71.
- Penedo, F.J., Dahn, J.R. (2005). Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Current opinion in psychiatry*, **18**(2), pp. 189–193.
- Petrillo, G., Caso, D., Donizzetti, A.R. (2004). Questionario “Salute e benessere nel mondo giovanile”, in *Promuovere la salute nei contesti educativi: comportamenti salutari e benessere tra gli adolescenti*, eds. G. Petrillo, D. Caso, Franco Angeli, Milano, (IT), pp. 401–407.
- Reckase, M. (2009). *Multidimensional Item Response Theory*. Springer, New York, (NY).
- Ries, F., Hein, V., Pihu, M., Armenta, J.M.S. (2012). Self-identity as a component of the theory of planned behaviour in predicting physical activity. *European Physical Education Review*, **18**(3), pp. 322–334.
- Samejima, F. (2016). Graded response model, in *Handbook of Item Response Theory*, ed. W.J. van der Linden, Springer, New York, (NY), pp. 95–108.
- Stephan, Y., Boiche, J., Trouilloud, D., Deroche, T., Sarrazin, P. (2011). The relation between risk perceptions and physical activity among older adults: a prospective study. *Psychology & health*, **26**(7), pp. 887–897.
- Sun, G., Acheampong, R.A., Lin, H., Pun, V.C. (2015). Understanding walking behavior among university students using theory of planned behavior. *International journal of environmental research and public health*, **12**(11), pp. 13794–13806.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**(2), pp. 461–464.
- Tudor-Locke, C., Bassett, D.R. (2004). How many steps/day are enough?. *Sports medicine*, **34**(1), pp. 1–8.
- World Health Organization. (2018). Physical activity. <https://www.who.int/news-room/fact-sheets/detail/physical-activity>.
- Wing Kwan, M.Y., Bray, S.R., Martin Ginis, K.A. (2009). Predicting physical activity of first-year university students: An application of the theory of planned behavior. *Journal of American College Health*, **58**(1), pp. 45–55.