POLITECNICO DI TORINO Repository ISTITUZIONALE

Validating Italian General Ecological Behaviour Questionnaire of Travellers using Dichotomous Rasch Model

Original

Validating Italian General Ecological Behaviour Questionnaire of Travellers using Dichotomous Rasch Model / Kumawat, Pinky; Pronello, Cristina. - ELETTRONICO. - (2022). (Intervento presentato al convegno TRB 101st Annual Meeting tenutosi a WASHINGTON DC nel 9-13 January, 2022).

Availability: This version is available at: 11583/2953560.4 since: 2022-01-27T20:46:11Z

Publisher: Transportation Research Board TRB

Published DOI:

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

Validating Italian General Ecological Behaviour Questionnaire of Travellers using Dichotomous Rasch Model

- 3
- 4
- 4 5

6 Pinky Kumawat

- 7 Politecnico di Torino, Dipartimento Interateneo di Scienze, Progetto e Politiche del Territorio
- 8 Viale Mattioli, 39, 10125, Torino, Italy
- 9 Tel: +39-011-0905640; Fax: +39-011-0906450; Email: pinky.kumawat@polito.it
- 10

11 Cristina Pronello, corresponding author

- 12 Politecnico di Torino, Dipartimento Interateneo di Scienze, Progetto e Politiche del Territorio
- 13 Viale Mattioli, 39, 10125, Torino, Italy
- 14 Tel: +39-011-0905613; Fax: +39-011-0906450; Email: cristina.pronello@polito.it
- 15
- 16
- Word count: 6,572 words text + 2 tables x 250 words (each) = 7,072 words
- 18
- 19
- 20
- 21 Submission Date: 26 July 2021

1 ABSTRACT

2 Ecological behaviour and the impact on environment are subjects of public concern and 3 understanding individual behavioural measures to induce sustainable lifestyles is of extreme 4 importance for policy makers to assess and promote sustainable mobility. To this end, a 5 questionnaire with highly reliable items, evaluation of determinants and accurate measurement of 6 ecological behaviour is a precondition for understanding the levers for a behavioural change. This 7 paper aims at understanding whether the dichotomous Rasch model provides a legitimate 8 measurement of General Ecological Behaviour (GEB) using a 26 items questionnaire as a valid 9 tool to assess pro-environment behaviour of a large sample of users. A web questionnaire was 10 administered using the snowball sampling plan in the Piedmont region (Italy) reaching out a sample of 4473 respondents. Results suggest that using the dichotomous Rasch model, proposed 11 12 questionnaire is able to effectively measure pro-environment behaviour of travellers. 13 Unidimensionality, perfect level of item reliability of 1, very high item separation of 34.22, 14 absence of larger differential item functioning, and local independence are all good indicators of a valid model. This research shows how a good, validated, and reliable measurement of ecological 15 16 behaviour would support public bodies to plan environment focused transport policies thanks to the knowledge of which variables determine the pro-environment behaviour. In addition, the 17 proposed approach allows also to measure the efficacy of the adopted policies. 18

19

20 Keywords: General Ecological Behaviour, Pro-environment Travel Behaviour, Dichotomous

21 Rasch Model

1 INTRODUCTION

2 Ecological behaviour and the impact of human activities on the natural environment are subjects 3 of public concern and have been largely studied in the psychological research that underlined the 4 importance of adopting more ecological behaviours or lifestyles (1, 2). The ecological behaviour 5 means the actions which contribute towards environmental preservation and conservation (3, 4). 6 It seems, albeit, that what people choose to do to reduce their environmental impact often does not 7 correspond well with what research suggests they should do (5, 6). This apparent lack of 8 correspondence has called into question the criterion validity of behavioural measures of 9 ecological lifestyles (7, 8). In this regard, the proper measurement of General Ecological 10 Behaviour (GEB) of users can serve as a powerful tool for policy makers to implement and, particularly, to assess more user-focused policies supporting people in adopting daily ecological 11 12 habits. For that a well-designed GEB questionnaire with proper items, that match the real lifestyle 13 habits of users is also a precondition and require attention, considering different cultural and 14 geographical contexts.

15 Therefore, various studies in literature used GEB to assess sustainable behaviour. Arnold 16 et al. (9) assessed electricity consumption of German adults; Kaiser and Wilson (10) used sample of two transport associations: one aims to promote a transport system that has as little negative 17 impact on humans and nature, the other represents automobile drivers' interests, such as proper 18 19 road maintenance, allowing higher speed limits on freeways, and fighting gasoline-tax increases. 20 Hergesell (11) examined differences in choosing the transport mode during the holidays through 21 general level of environmental commitment across lifestyle domains and found that train users 22 tend to be more environmentally committed compared to car users. Two versions of GEB 23 questionnaire were proposed to assess pro-environment travel behaviour in an Italian region. A 24 first version was proposed by Gaborieau and Pronello (12) based on Kaiser and Wilson (10), 25 which we call GEB-40 (40 dichotomous items); the second version was proposed by (13) as an 26 extended version of GEB-40, which we call GEB-51 (51 dichotomous items). One of the 27 weaknesses of previous two Italian GEB versions (GEB-40 and GEB-51) was the inclusion of 28 irrelevant and redundant items that were excluded in this study.

29 At best of our knowledge, the studies measuring GEB questionnaire using the Rasch model 30 (14), whether in different cultural contexts or in a single area, used limited and small sample size. Kaiser and Biel (15) compared ecological behaviour of 247 Swedish and 445 Swiss people; Kaiser 31 and Wilson (10) compared 686 Californian students and 445 Swiss participants; Gaborieau and 32 33 Pronello (12) compare 131 Italian, 445 Swiss, and 247 Swedish participants; Hergesell (16) based 34 on a sample of 349 German citizens, although the sample size is still within acceptable boundaries, 35 according to Linacre (17). Nevertheless, replication in a larger sample of population is highly desirable and the use of small samples was reported as one of the limitations of previous researches 36 37 (9, 12).

The current research focused to obtain high item reliability, good separation indexes, and well-functioning items with a larger sample size. In addition, to reduce the fatigue of respondents, it has been paid attention to use comparatively few (26) and highly reliable items to assess the GEB. The paper has three main objectives:

42

• To determine whether the dichotomous Rasch model could provide a legitimate measure of the chosen 26 items in the polytomous GEB questionnaire as a valid tool to assess the proenvironment behaviour of users in Piedmont region, Italy; To check the validity of dichotomous scale measurement instead of original polytomous
 questionnaire, with a larger sample size, to allow a comparison with the previous two versions of
 GEB questionnaires (GEB-40 and GEB-51) in the Italian context;

• To understand if the obtained GEB Rasch person measure has some impact on travel behaviour – mode choice – to assess if the people behaving more ecologically effectively chose sustainable modes or people behaving less ecologically chose unsustainable transport modes.

7

8 The paper is organised as follows: the following section will present the methodology used 9 to design and administer the questionnaire, the sampling plan, and the requirements to assess the 10 dichotomous Rasch model. Section 3 presents the obtained results. Then, section 4 discusses the 11 appropriateness of the dichotomous scale and questionnaire items, the inclusion or exclusion of 12 items, and some aspects related to questionnaire design. Finally, the discussion and conclusions 13 are presented.

13

15 **METHODOLOGY**

The research was conducted in the Piedmont region (Italy), with focus on the metropolitan area of Torino. A web questionnaire has been designed to get in depth information related to opinions, preferences, attitudes, lifestyles, and mobility patterns of users with the aim of studying the proenvironmental behaviour of the sample and understanding whether a general pro-environment attitude may legitimately be assessed using the Rasch model. A four-step methodology comprised: (1) survey design; (2) survey administration and sample selection; (3) data base construction; (4)

22 model estimation and testing of GEB.

23

24 Survey Design

25 A survey has been designed, named "Come ci muoviamo? ... ma soprattutto come vorremmo 26 muoverci?". The survey is composed by two different web-questionnaires. The first part includes 27 questions well established in literature, which can ensure well-grounded comparison, and it is 28 composed by six sections: mobility in a standard week; travel diary related to the most important 29 trip; integrated mobility; Mobility as a Service; attitudes and preferences – including GEB; and 30 socio-economic data. The second part is composed by new questions, derived from recent results 31 from behavioural theories to overcome some gaps observed in previous researches by (12, 13) and 32 it is composed by two sections: information about the most important trip; and attitudes and 33 preferences related to this trip. This paper mainly focuses on analysing the general attitudes 34 towards the environment and its ecological behaviour using the section of the questionnaire related 35 to GEB.

The GEB questionnaire is based on GEB-40 and GEB-51 but includes only 26 items (GEB-26) reported in Table 1, resulting from deleting redundant and problematic items found in GEB-40 and GEB-51. The questionnaire has been designed to collect polytomous data based on a 6-point Likert scale where 1 was "completely disagree" and 6 "completely agree".

- 40
- 41 42

TABLE 1 Structure of GEB-26 Questionnaire

No.	Item description					
	Category 1 - Pro-social behaviour					
1	Sometimes I give money to panhandlers	CS1				
2	From time to time, I give money to charity	CS2				
3	If an elderly or disabled person enters a crowded PT vehicle, I offer him/her my seat	CS3				

4	If I were an employer, I would not hesitate to hire a person previously convicted of	CS4						
5	Crime Sometimes Uride public transport without paying a fare							
5	Category 2 - Ecological garbage handling							
6	I put dead batteries in the garbage	R1(-)						
7	I sort glass wastes for recycling	R5						
-	Category 3 - Water and power saving							
8	I turn off the heat at night	AE4						
9	I wait until I have a full load before doing my laundry	AE5						
10	In winter, I leave the windows wide open for long periods of time to let in fresh air	AE6(-)						
	Category 4 - Ecologically aware consumerism							
11	I use fabric softener with my laundry	CE1(-)						
12	I always look to buy vegetables from biological agriculture	CE6						
13	Sometimes, I sell goods I don't use anymore	CE7						
14	Sometimes, I buy second hand goods	CE8						
15	Sometimes, I offer goods I don't use anymore	CE9						
16	Sometimes, I rent for free to someone, goods I occasionally use	CE14						
17	I eat less meat than years ago	CE15						
	Category 5 - Garbage inhibition							
18	I re-use plastic bag from the groceries	RR1						
19	I sometimes buy beverage in cans	RR2(-)						
	Category 6 - Environmental activism							
20	I often talk with friends about problems related to the environment	V1						
21	I am a member of an environmental organization	V2						
22	In the past, I have pointed out to someone his or her un-ecological behaviour	V3						
23	I sometimes contribute financially to environmental organizations	V4						
24	I boycott companies using OGM or pesticides	V5						
	Category 7 - Transport							
25	Usually, I do not drive my automobile in the city	T1						
26	I usually drive on freeways at speeds lower than 100km/h	T2						

Note: (-) items positively formulated as environmentally damaging, recoded.

1 2

3 Survey Administration and Sample Selection

4 The survey was administered to the population living in the Piedmont region, with focus on 5 metropolitan area of Torino. The citizens were reached through different channels: email, flyers, 6 notice on the websites of municipalities and transport companies, formal notice to employees in 7 Rail Infrastructure Managers, direct contact with major cultural and sport associations, 8 newspapers, and local radio and Twitter including the survey in traffic bulletin. The link to the 9 survey and QR code were available through the above channels and respondents filled in the 10 questionnaire using the Computer Assisted Web Interviewing (CAWI), developed using the 11 software Lime Survey.

Such wide dissemination was possible thanks to the support from the Local Public Bodies – Piedmont Region, City of Torino, main universities (Politecnico di Torino and Università degli Studi di Torino), the transport authority Agenzia Mobilita Piemontese and some transport operators as Gruppo Torinese Transporti and Sadem and the Rete Ferroviaria Italiana. Answers were collected in the period from the 27th of October 2017 to the 24th of April 2018, based on the snowball sampling plan, reaching out a random sample of 4473 respondents.

- 18
- 19 Database Construction

1 The initial sample of 4473 records was resized to 4212 units excluding the persons whose 2 destination was outside both Italy and the region. The 4212 records have been used in Rasch model 3 estimation. The residential locations are classified in three areas, urban (metropolitan area of 4 Torino), suburban (municipalities around Torino-first belt) and rural (rest of the territory-second 5 belt). The Piedmont Territorial Demographic Observatory identifies a "first" and a "second" belt 6 of municipalities surrounding Torino¹. The majority of respondents come from urban area and the 7 distribution of the three residential locations is: 2154 (51.14%) urban, 740 (17.57%) suburban, and 8 1318 (31.29%) rural.

9 The next step for constructing the data base was the check of missing values. Two 10 variables, T1 and T2 related to category 7 "transport" (Table 1), have, respectively, 437 and 572 not applicable responses. These are not missing at random, but they are a choice from respondents, 11 12 and they were considered as missing during the analysis to avoid any imputation, having a large 13 database. The software Winsteps, used for the Rasch model, does not require complete data in 14 order to make estimates, because it uses Joint Maximum Likelihood Estimation (JMLE) that is 15 very flexible as regards estimable data structures. Waterbury (18) reports that Rasch model can 16 handle varying amounts of missing data, provided that the missing responses are not missing not at random. Hence, the missing records without any imputation were used whereas other variables 17 have complete data for corresponding records. Finally, the dataset is converted from polytomous 18 19 scale to dichotomous scale by converting the first three categories, from 1 (completely disagree) 20 to 3, to 1 "No"; and second three categories, from 4 to 6 (completely agree), to 2 "Yes".

21

22 Rasch Model as a Measure of General Ecological Behaviour

23 The general attitude towards the environment, based on the data collected by the GEB 24 questionnaire, was analysed using Rasch Model for scale measurement. Rasch analysis describes 25 procedures that use a particular model with outstanding mathematical properties developed by 26 Georg Rasch (14) for the analysis of data from tests and questionnaires. The mathematical theory 27 underlying Rasch models is a special case of Item Response Theory (IRT) and, more generally, a 28 special case of a generalized linear model. The statistical calculations employed by the Rasch 29 model to locate and order persons and item difficulty is based on Guttmann Scaling and it can be 30 used with both dichotomous and polytomous data sets (19). This study explores the potential of 31 using the dichotomous Rasch model to analyse polytomous items for GEB attitude measure.

The Dichotomous Rasch model (DRM) (14) is the simplest model in the Rasch family of models. It was designed for use with ordinal data that are scored in two categories. The DRM uses sum scores from these ordinal responses to calculate interval-level estimates that represent person locations and item locations on a linear scale that represents the latent variable. The difference between person and item locations can be used to calculate the probability for a correct or positive response (x = 1), rather than an incorrect or negative response (x = 0). The equation for the DRM is reported in eq. 1:

39

40
$$B_n - D_i = \ln(P_{ni}/P_{ni})$$

(1)

41

43

42 where

 B_n = ability of a specific person n;

¹https://web.archive.org/web/20140727134854/http://www.demos.piemonte.it/site/images/stories/caricafil e/territori/E_area_metropolitana.pdf.

1 D_i = difficulty of a specific item i; 2 P_{ni} = probability of person n correctly answering item i; and 3 ln = "log-odds units" (logits), which is a natural logarithm. 4 The DRM specifies the probability, P, that the person n with ability B_n succeeds on item i 5 of difficulty D_i. 6 The key Rasch model requirements are unidimensionality, local independence, persons-7 invariant item estimates/person parameter separability, and item-invariant person estimates/item 8 parameter separability. 9 For the parameter estimation for DRM, Winsteps Rasch Analysis program was used. 10 Winsteps implements two methods of estimating Rasch parameters from ordered qualitative observations: JMLE also known as UCON (Unconditional Maximum Likelihood Estimation) (20) 11 12 and PROX (Normal Approximation Algorithm) devised by Cohen (21). 13 14 Rasch Measures and Model fit 15 Rasch model fits are used to examine the unidimensionality of the latent trait to measure attitude 16 towards GEB. Unidimentionality is evaluated using: 1) point-biserial correlation 2) fit statistics, 17 3) Principal Component Analysis of Residuals, and 4) local independence. 18 19 **Point-biserial Correlation** Point-biserial correlation is a useful diagnostic indicator of data 20 miscoding or item mis-keying: negative or zero values indicate items or persons with response 21 strings that contradict the variable. Li et al. (22) suggests that point-measure correlation larger than 22 .3 indicate that items are measuring the same construct. 23 24 Fit Statistics Rasch model provides two indicators of misfit: INFIT and OUTFIT. Since the ZSTD 25 value is based on the MNSQ, as reported by Boone et al. (23), we first examine the MNSQ for 26 evaluating fit. If the MNSQ value lies within an acceptable range, we ignore the ZSTD value. 27 According to Boone et al. (23), INFIT and OUTFIT mean-square fit statistics between 0.5-1.5 28 represents productive items. For mathematical formulation of point-biserial correlation, INFIT, 29 OUTFIT, and ZSTD refer to (12). 30 31 Principle Component Analysis of Residuals (PCAR) Unidimensionality was checked through 32 PCAR. According to Reckase (24) unidimensionality is hold if: a) the amount of variance 33 explained by measures is > 20%; b) unexplained variance of the eigenvalue for the first contrast is 34 < 3; and unexplained variance accounted by first contrast is < 5%. 35 36 **Local Independence** Local independence means that after the contribution of the latent trait(s) to 37 the data is removed, all that is left is random noise (25). A correlation of r=0.40 among items is 38 low dependency. 39 Besides these, Rasch model assumptions include assessing reliability and separation of 40 measures, differential item functioning, evaluation of item difficulty using Write map to evaluate 41 construct validity. 42 43 *Reliability* It ranges from 0 to 1 and the higher is better (26). Bond and Fox (27) suggested value 44 between 0.6-0.8 is acceptable. 45

Separation index Separation index of 1.50 represents an acceptable level, 2 represent a good level
 according to Miller and Dishon (28) and 3 represents an excellent level as reported by Duncan et
 al. (29).

4

5 Differential Item Functioning (DIF) DIF is used to determine whether the individual items on a 6 test function in the same way for two or more groups (30). Mantel-Haenszel (MH) (31) test for 7 dichotomies is used. Items are flagged as DIF when the MH probability value is ≤ 0.05 and then 8 the DIF size is assessed according the criteria by Zwick et al. (32). Moderate to large DIF when 9 size CUMLOR is ≥ 0.64 , slight to moderate DIF when size CUMLOR is ≥ 0.43 , negligible when 10 size CUMLOR < 0.43. We investigated DIF by two criteria: 1) gender and 2) residential location.

1112 **RESULTS**

13 This section presents the results by following the various steps described in the methodology. 14

15 **Point-biserial Correlations**

All items' correlations are positive and pointing in the same direction. However, three smallpositive correlations are observed and analysed hereafter:

18

Item AE6_REVC has a low correlation (.05) close to zero. When assessing closely this item,
 74.17% users agree, and 25.83% disagree, showing that this is one of the easiest behaviours to
 engage into (Measure=-0.76);

• *Item CS6_REVC* has a low correlation (.09) close to 0.1. When closely assessing this item, 90.38% users agree, and 9.62% disagree; similar to the previous item, this is also one of the easiest behaviours to engage into (Measure=-2.08). Almost most of the users agreed to using public transport without tickets, which may cause the low correlation;

Item CS4 has a low correlation (.08) near to 0.1. Analysing this, no big difference among
 the answered categories of the respondents (46.77% disagree and 53.23% agree) was found. This
 item seems to have medium difficulty across all respondents (Measure = 0.31).

30 Fit statistics

31 Item AE6_REVC has the highest mean-square outfit (1.55). The small difference of .05 over the 32 threshold might not degrade the measurement. We found that all other items are within acceptable 33 ranges of MNSQ, hence we are not investigating ZSTD.

34

35 Principle Component Analysis of Residuals

36 First, the amount of variance explained by measures is 34.2% (11.5% of raw variance explained

37 by persons and 22.7% of raw variance explained by items) which is larger than the requirement of

38 20% according to Reckase (24). Second, the unexplained variance by first contrast is 5.4%, which

is slightly greater than 5%, but the eigenvalue of first contrast is 2.14 (< 3). The results suggest that the unidimensionality is hold across the whole test.

40 that the undimensionality is hold across the whole test. 41 The loading of items on the first contrast of the residual based PCA are shown in Figure

41 The loading of items of the first contrast of the residual based FCA are shown in Figure 42 1, showing that this possible sub-dimension is formed by two items, A (AE6_REVC), and B

- 43 (CS6_REVC). Items A and B have the largest loadings, quite far away from the general cluster
- 44 created by the other items, and the eigenvalue of first contrast is 2.14 (~2 items). To see the items
- 45 corresponding to the letter of alphabet represented in Figure 1, refer to Table 2.

Entry	Total	Measure	Model	Infi	t	Out	fit	Point-bis	. Corr.	Exact M	atch (%)	Item
No.	Score		S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%	
10	7336	76	.04	1.24	9.90	1.55	9.90	A .05	.33	70.9	75.7	AE6_REVC
5	8019	-2.08	.05	1.10	2.54	1.48	6.48	B .09	.24	90.4	90.4	CS6_REVC
4	6454	.31	.03	1.27	9.90	1.43	9.90	C .08	.38	53.0	65.8	CS4
6	7938	-1.86	.05	1.08	2.17	1.31	4.87	D .15	.26	88.0	88.5	R1_REVC
19	6418	.35	.03	1.20	9.90	1.29	9.90	E .16	.38	56.6	65.8	RR2_REVC
25	5712	.46	.04	1.15	9.90	1.24	9.90	F .20	.37	58.6	65.5	T1
11	6248	.53	.03	1.14	9.90	1.22	9.90	G .23	.38	59.8	65.8	CE1_REVC
26	5285	.78	.04	1.09	7.13	1.12	6.22	Н .28	.38	61.4	66.6	T2
8	7203	58	.04	.99	72	.99	26	I .35	.34	73.9	73.3	AE4
13	5535	1.37	.04	.97	-1.91	.95	-2.17	J .40	.37	72.7	72.5	CE7
1	5491	1.43	.04	.96	-2.25	.94	-2.33	K .41	.37	74.1	73.1	CS1
14	5812	1.03	.03	.96	-3.36	.93	-3.66	L .42	.38	70.1	68.8	CE8
18	8185	-2.69	.07	.94	-1.15	.87	-1.55	M .26	.19	94.4	94.3	RR1
2	5949	.87	.03	.93	-5.96	.91	-5.22	m .45	.38	70.8	67.5	CS2
7	8176	-2.64	.07	.93	-1.22	.76	-2.93	1 .27	.20	94.1	94.1	R5
3	8136	-2.48	.06	.92	-1.62	.76	-3.29	k .30	.21	93.2	93.2	CS3
9	7985	-1.98	.05	.92	-2.03	.85	-2.45	j .32	.25	90.2	89.6	AE5
16	6673	.06	.03	.91	-7.93	.88	-7.00	i .47	.37	70.6	66.8	CE14
17	6441	.32	.03	.91	-8.75	.86	-8.55	h .48	.38	70.7	65.8	CE15
12	5911	.92	.03	.90	-8.18	.87	-7.59	g .48	.38	72.5	67.8	CE6
21	4586	3.13	.06	.90	-2.48	.72	-4.61	f .38	.27	91.5	91.3	V2
22	6715	.01	.03	.90	-8.58	.87	-7.08	e .47	.37	72.2	67.2	V3
15	7134	49	.04	.89	-7.71	.83	-6.86	d .47	.35	76.3	72.2	CE9
20	6625	.11	.03	.88	-9.90	.83	-9.90	c .50	.37	72.4	66.6	V1
24	5391	1.56	.04	.88	-6.88	.83	-6.85	b .48	.36	78.2	74.8	V5
23	4912	2.33	.04	.83	-6.48	.66	-8.91	a .50	.32	84.9	84.0	V4
Mean	6548.8	.00	.04	.99	-1.0	1.00	9	-	-	75.4	75.7	-
P.SD	1059.6	1.50	.01	.12	6.	.24	6.7	-	-	12.0	10.6	-

TABLE 2 Estimates of Item Parameters, Infit, Outfit, and Point-biserial Correlations

 STANDARDIZED RESIDUAL CONTRAST 1 PLOT



FIGURE 1 Item loadings on the first contrast

5 The correlations of the person measures computed with each cluster of items were as 6 follows: Cluster 1 and 2: r = 1.0; Cluster 1 and 3: r = 0.0587; Cluster 2 and 3: r = 1.0. Having 7 cluster 1 and 3 low correlation, the sub-dimension might be due to the items in cluster 1, as 8 discussed above for items A and B.

9 10 Local independence

11 According to the Linacre guidelines² all items correlation is <0.4, hence no item residuals are 12 correlated, respecting the local independence assumptions of Rasch analysis.

13

1 2 3

4

Person measure reliability is .67 and *item measure reliability* is 1 (perfect), acceptable with the less variability of the measurement attributed to measurement error. The *person separation*, equal to 1.44, shows that this test can distinguish between high and low performers (1.44, ~2 levels) and represent good level of separation according to Miller and Dishon (28).

- 18
- 19 *Item separation* is very high, equal to 34.22, and represent excellent level of separation (29). With
- 20 this large person sample, the item difficulties are estimated exceedingly precisely and validating

21 the GEB construct validity (>3).

22

² <u>https://www.winsteps.com/winman/table23_99.htm.</u>

1 *DIF* is assessed using MH test, based on gender (Female, Male) and residential location (Urban,

Suburban, Rural), and it is conducted by comparing a reference group (the majority group) with a
focal group (the minority group) (*33*). The reference group for gender and for residential location
are respectively, female and urban while the other are the focal groups.

5 Considering gender, two items report DIF of slight to moderate size: CE9 with p value .00 6 and DIF size .63; and V1 with p value .00 and DIF Size -.47. Looking at residential location, two 7 items, R5 and T1, show moderate to large DIF. R5 with p value .00 and DIF size .90 for urban and 8 rural; p value .00 and DIF size 1.12 for urban and suburban. T1 with p value .00 and DIF Size .44.

9 10 Write map

Figure 2 depicts the person measures (left) and the item measures (right). Persons at the top had the least difficulty endorsing items, while persons at the very bottom had the most difficulty endorsing items. We can observe that:

14

• The most difficult item is V2 followed by item V4; both belong to the category of environmental activism;

• The easiest items are R5 and RR1, followed by CS3. These three items are not targeting to any person; some persons above and below these items are less inclined to GEB, so these items are not useful to the GEB measurement but still fall within the user's ability range;

Items CS1 and V5 measure similar portions of the trait and therefore, from a measurement
 perspective, are redundant. This is also the case of items CE6, CS2, T2; CE1_REVC, T1; CE15,
 CS4, RR2_REVC; CE14, V3; AE4, CE9; AE5, CS6_REVC; and R5, RR1. Within groups of items,
 individual items can be removed losing a small precision of the measurement;

• We do not see the gap between items more than a logit, but there is a need to fill the measurement gaps, between V4 and CS1, and between items AE6_REVC and R1_REVC. This explains the relatively poor value of the individual separation reliability.

28 DISCUSSION AND CONCLUSIONS

The purpose of this research was to scrutinize psychometric properties of the GEB-26 questionnaire using a DRM approach to validate and compare the scale with those used in previous research and to understand if this has some impact on travel behaviour, notably on mode choice.

32 Unidimensionality has been evaluated utilizing Rasch fit statistics, as well as PCAR and 33 point-biserial correlations. Notably, all these tests of the measure's dimensionality suggest the 34 items lie on one trait as hypothesized during survey design. Therefore, it can be recommended to 35 use the GEB-26 as a unidimensional scale. Model fit indicators suggest that the scale contains one particularly misfitting item, AE6_REVC, with only slightly high outfit MNSQ value (0.05), that 36 37 is not threatening the validity of the scale, so that it is not suggested to delete it. The fact that item 38 AE6 REVC was the only item with poor fit warrants further investigations as it offers potential 39 insights into the structure of GEB. It is well known that negatively coded items, especially if there 40 are only a few and located at the end of the questionnaire, may be confusing for the respondents 41 (34). However, it is also possible that the item did not confuse the respondents, but not behaving ecologically may actually not be seen as an inverse conceptualization of ecological behaviour, but 42 43 rather a (partly) different construct in its own right. Moreover, local independence, reliability, and 44 separation indexes assumptions were confirmed with good Rasch measures validity.

45



FIGURE 2 Write map

We have obtained perfect level of reliability of 1, separation of 34.22 for items, and 6 sufficient level of person separation and reliability. Although, person (test) reliability mainly 7 depends on the variance of sample ability and on the number of categories per item. If we have 8 more categories, then we might achieve higher person reliability. So, in this study we first validated 9 the questionnaire by converting polytomous scale to dichotomous scale to compare the results 10 from the previous studies (GEB-40 and GEB-51) and to verify how the selected test performs with 11 larger sample size as also person separation and reliability are sample dependent. The most 12 important aspect is to validate the questionnaire items that have been selected, to revise them, if 13 necessary, for designing the next survey.

Observing DIF analysis, item CE9 is more difficult for females and V1 is more difficult for males. This shows the cultural, societal, and attitudinal difference as determinant factors to engage in a certain behaviour. The DIF size for these two items was slight to moderate, hence we are not considering excluding these items for the next questionnaire. This aspect is also part of

1 Campbell's paradigm (35) of attitude, which states that some behaviour may be more difficult in 2 certain contexts than in others. This applies also to the residential location (R5) and the related 3 land use; the results show how a well shared habit of sorting glass for recycling is easier for people 4 living in rural areas due to the different organisation of collection points of glass located at single 5 homes, differently from the scattered patterns of collection points in the cities. The way of living 6 in rural areas makes also people less used to drive in congested urban traffic (T1), reason why 7 urban citizens are more used and, thus, inclined to use car to travel inside the cities; differently, 8 those living outside prefer travelling to the city by train or suburban bus to avoid traffic and parking 9 problems. So, what stated by Arnold et al. (9) holds true, showing the importance of surroundings 10 and contextual elements in the daily routine. The DIF size for R5 is moderate to large, which requires some attention to consider in further analysis; instead, item T1 has slight to moderate DIF, 11 12 not necessarily indicated for deletion.

13 Observing the results, GEB-26 shows good psychometric properties when using DRM to 14 validate the scale; some further analysis can be useful to verify the three items that are slightly 15 borderline.

16 The second aspect that was investigated, concerning the validity of GEB in influencing the modal choice, is key in the current debate on climate change that calls for major changes in 17 people's daily lifestyles (2). A frequent question arising is: do what people report to protect the 18 19 environment converge with their environmental impact? If, theoretically speaking, this could hold 20 true, under an empirical observation our results show the opposite. We observed that out of selected sample of 4212 respondents, for the most important trip (that with longest distance), 1368 21 22 (32.48%) use trip chain followed by 1156 (27.45%) using car, 729 (17.31%) using public transport, 23 330 (7.83%) walking, and 310 (7.36%) cycling. Looking into trip chain, car as driver is used by 24 the highest percentage of respondents, 1333 (31.65%), followed by 1096 people (26.02%) using 25 public transport, 667 travelling by train, 401 walking, and 322 cycling. This finding shows how people do not do what they intent/say to do. Hence, behavioural measures of ecological lifestyles 26 27 may reflect actual environmental impact in some other contexts such as in electricity consumption, 28 as reported by Arnold et al. (9), but they do not apply in transport sector by looking at results and 29 as shown in previous studies (36). This is referred as attitude-behaviour gap (37) or behaviour 30 intention gap (38), demonstrating the volatility of the concepts of attitude or intention (39). The results obtained in this research also contradict what found in (12), where the high GEB score was 31 32 correspondent to those users who use soft modes (walking or bike) for their most frequent trip, 33 followed by public transport (regional train, bus, tram or metro) and, then, private motorised 34 vehicles (car or motorbike). One reason of this contradiction might be that the trip chain was 35 excluded by Gaborieau and Pronello (12) and the sample was smaller (108 users). This discrepancy 36 will be further investigated in the continuation of the research.

37 It should also be recalled that the sample sizes in previous studies – in Italian context (GEB-38 40 and GEB-51), in Swedish and Swiss context (15), and in Californian context (10) – were 39 comparatively too small, although still within acceptable boundaries, according to Linacre (17). 40 Nevertheless, replication in a larger sample is highly desirable as suggested in current research. Regarding the generalizability of the results, it must be noted that the composition of samples of 41 previous studies was formed thanks to a stratified sampling plan. Thus, different results may be 42 observed when the sample follows the snowball sampling approach and the participants are, as in 43 44 this case, younger and/or with a bit lower educational level. Finally, it needs to be emphasized that 45 even excellent internal validity is no assurance that a given scale will also exert good external 46 validity.

One of the limitations of these studies assessing ecological and environmental behaviour is that people may not be aware about their environmental impacts and/or the damage they cause to the environment. As reported by Hamidi and Zhao (40), the individuals who have greater environmental awareness are more likely to travel by public transport or cycling if the physical conditions facilitate using these modes. Hence, the proper environmental and mobility education is needed to educate people as also suggested by Gaborieau and Pronello (12); and Pronello and Camusso (36).

8 Further research is needed to deepen our understanding of the GEB and to devise 9 appropriate measurement instruments. There was no evidence that individuals with diverging 10 sociodemographic characteristics, such as age, had a different understanding of the items. The item which is difficult could be answered by respondents with high capability, whilst easy items could 11 12 be answered by respondents with high and low ability. Overlapping items measure different 13 elements with different levels of difficulty (41), hence we do not suggest to exclude items to design 14 a new survey by looking only at redundancy of items in write map. Some recommendations are 15 worthy to be given for improving the scale. Firstly, more items could be selected with high or low 16 difficulties so that the scale will be able to measure individuals outside an intermediate level of ecological behaviour, particularly to fill in the gaps identified in the study in write map analysis. 17 This is important because limited differentiation capabilities may attenuate existing effects of 18 19 measuring ecological behaviour. The GEB-26 might not be capable of detecting strong effects 20 potentially attributable to interventions based on ecological behaviour in terms of larger person 21 ability range due to weakness of questionnaire design; in fact, we obtained person measure 22 reliability equal to .67 and person separation equal to 1.44, which is acceptable but not excellent. 23 Hence, GEB researchers would profit from more sensitive measurement instruments capable of 24 detecting differences between individuals who are high and low in ecological behaviour. 25 Furthermore, we do not suggest excluding any item by looking only at the dichotomous scale 26 measurement. The item exclusion will be further decided after measuring the original 6 scale 27 polytomous questionnaire using Rasch rating scale model, which is the next step of our research, continuing to validate and select the appropriate measurement scale to measure GEB of users. As 28 29 suggested by Linacre³, the scale with more categories is expected to give better and higher person 30 reliability and separation. Future research may also investigate by testing the GEB questionnaire in different cultural and territorial contexts such as different regions, cities, and metropolitan areas 31 32 of Italy, and different European countries to validate the appropriate GEB questionnaire.

In summary, we can conclude that GEB-26 shows acceptable approximation to Rasch requirements. Improvements, as outlined above, are strongly warranted, and may yield a reliable and internally valid measurement device for the measurement of GEB.

36 The final aim is to propose proper public targeted policies to induce people to sustainable 37 travel choices. For that, a well-planned and environment-focused transport education policy can 38 play a role to educate people and make them aware about their environmental footprints and 39 motivate them to behave ecologically and sustainably. In this regard, the transport policies together 40 with the idea of giving incentives to people when they use the sustainable modes could trigger them towards more sustainable behaviour as reported by Pronello and Kumawat (42). Technology 41 can also play a role for giving incentives to promote sustainable mobility or engaging them in pro-42 43 environment ecological behaviour with the help of smartphone apps, as these days the apps are 44 becoming part of daily life of people and the trend is exponentially increasing (42). To this end, a

³ https://www.winsteps.com/winman/reliability.htm.

1 good, validated, and reliable measurement of ecological behaviour would let public bodies to

- 2 measure the efficacy of adopted policies.
- 3

4 AUTHOR CONTRIBUTION STATEMENT

- 5 The authors confirm contribution to the paper as follows: study conception and design: Cristina
- 6 Pronello; data collection: Cristina Pronello; analysis of results: Pinky Kumawat; interpretation of
- 7 results: Pinky Kumawat and Cristina Pronello; draft manuscript preparation: Pinky Kumawat and
- 8 Cristina Pronello.
- 9 All authors reviewed the results and approved the final version of the manuscript.
- 10

1 **REFERENCES**

- Howard, G. S. Adapting Human Lifestyles for the 21st Century. *American Psychologist*,
 Vol. 55, No. 5, 2000, pp. 509–515. https://doi.org/10.1037/0003-066X.55.5.509.
- Otto, S., F. G. Kaiser, and O. Arnold. The Critical Challenge of Climate Change for
 Psychology: Preventing Rebound and Promoting More Individual Irrationality. *European Psychologist*, Vol. 19, No. 2, 2014, pp. 96–106. https://doi.org/10.1027/1016 9040/a000182.
- Axelrod, L. J., and D. R. Lehman. Responding to Environmental Concerns: What Factors
 Guide Individual Action? *Journal of Environmental Psychology*, Vol. 13, No. 2, 1993, pp.
 10 149–159. https://doi.org/10.1016/S0272-4944(05)80147-1.
- Maloney, M. P., and M. P. Ward. Ecology: Let's Hear from the People: An Objective Scale
 for the Measurement of Ecological Attitudes and Knowledge. *American Psychologist*, Vol.
 28, No. 7, 1973, pp. 583–586. https://doi.org/10.1037/h0034936.
- 14 5. C Stern, P. Toward a Coherent Theory of Environmentally Significant Behaviour. *Journal* 15 of Social Issues, Vol. 56, No. 3, 2000, pp. 407–424.
- Gatersleben, B., L. Steg, and C. Vlek. Measurement and Determinants of Environmentally
 Significant Consumer Behavior. *Environment and Behavior*, Vol. 34, No. 3, 2002, pp. 335–
 362. https://doi.org/10.1177/0013916502034003004.
- Gifford, R., C. Kormos, and A. McIntyre. Behavioral Dimensions of Climate Change:
 Drivers, Responses, Barriers, and Interventions. *WIREs Climate Change*, Vol. 2, No. 6,
 2011, pp. 801–827. https://doi.org/10.1002/wcc.143.
- Steg, L., and C. Vlek. Encouraging Pro-Environmental Behaviour: An Integrative Review
 and Research Agenda. *Journal of Environmental Psychology*, Vol. 29, No. 3, 2009, pp.
 309–317. https://doi.org/10.1016/j.jenvp.2008.10.004.
- Arnold, O., A. Kibbe, T. Hartig, and F. G. Kaiser. Capturing the Environmental Impact of Individual Lifestyles: Evidence of the Criterion Validity of the General Ecological Behavior Scale. *Environment and Behavior*, Vol. 50, No. 3, 2018, pp. 350–372. https://doi.org/10.1177/0013916517701796.
- Kaiser, F. G., and M. Wilson. Assessing People's General Ecological Behavior: A CrossCultural Measure1. *Journal of Applied Social Psychology*, Vol. 30, No. 5, 2000, pp. 952–
 978. https://doi.org/10.1111/j.1559-1816.2000.tb02505.x.
- Hergesell, A. Environmental Commitment in Holiday Transport Mode Choice.
 International Journal of Culture, Tourism and Hospitality Research, Vol. 11, No. 1, 2017,
 pp. 67–80. https://doi.org/10.1108/IJCTHR-09-2015-0118.
- Gaborieau, J.-B., and C. Pronello. Validation of a Unidimensional and Probabilistic
 Measurement Scale for Pro-Environmental Behaviour by Travellers. *Transportation*, Vol.
 48, No. 2, 2021, pp. 555–593. https://doi.org/10.1007/s11116-019-10068-w.
- 38 13. Duboz, A. The Intention to Use Real-Time Multimodal Information to Change Travel
 39 Behaviour The Use of Psychosocial Variables for the Market. Politecnico di Torino, Italy,
 40 2018.
- Rasch, G. Probabilistic Models for Some Intelligence and Attainment Tests. *Information and Control*, Vol. 4, No. 4, 1961, p. 382. https://doi.org/10.1016/S0019-9958(61)80061-2.
- Kaiser, F. G., and A. Biel. Assessing General Ecological Behavior: A Cross-Cultural Comparison between Switzerland and Sweden. *European Journal of Psychological Assessment*, Vol. 16, No. 1, 2000, pp. 44–52. https://doi.org/10.1027//1015-5759.16.1.44.
- 46 16. Hergesell, A. Exploring the Relationship between General Environmental Behaviour and

- Holiday Transport Mode Choice. *BEST EN Think Tank XIV: Politics, Policy and Governance in Sustainable Tourism*, 2014, pp. 142-149.
- Linacre, J. M. Sample Size and Item Calibration or Person Measure Stability. *Rasch Measurement Transactions*, Vol. 7, No. 4, 1994, p. 328.
- 18. Waterbury, G. T. Missing Data and the Rasch Model: The Effects of Missing Data Mechanisms on Item Parameter Estimation. *Journal of applied measurement*, Vol. 20, No. 2, 2019, pp. 1–12.
- 8 19. Pallant, J. F., and A. Tennant. An Introduction to the Rasch Measurement Model: An 9 Example Using the Hospital Anxiety and Depression Scale (HADS). British Journal of 10 Clinical Psychology, Vol. 46, No. 1. 2007, 1 - 18.pp. https://doi.org/10.1348/014466506X96931. 11
- Wright, B., and N. Panchapakesan. A Procedure for Sample-Free Item Analysis.
 Educational and Psychological Measurement, Vol. 29, No. 1, 1969, pp. 23–48. https://doi.org/10.1177/001316446902900102.
- Cohen, L. Approximate Expressions for Parameter Estimates in the Rasch Model. *British Journal of Mathematical and Statistical Psychology*, Vol. 32, No. 1, 1979, pp. 113–120.
 https://doi.org/10.1111/j.2044-8317.1979.tb00756.x.
- Li, C.-Y., S. Romero, H. S. Bonilha, K. N. Simpson, A. N. Simpson, I. Hong, and C. A.
 Velozo. Linking Existing Instruments to Develop an Activity of Daily Living Item Bank. *Evaluation & the Health Professions*, Vol. 41, No. 1, 2018, pp. 25–43.
 https://doi.org/10.1177/0163278716676873.
- Boone, W. J., J. R. Staver, and M. S. Yale. *Rasch Analysis in the Human Sciences*. Springer
 Netherlands, 2014, pp. 1-482. https://doi.org/10.1007/978-94-007-6857-4.
- 24 24. Reckase, M. D. Unifactor Latent Trait Models Applied to Multifactor Tests: Results and
 25 Implications. *Journal of Educational Statistics*, Vol. 4, No. 3, 1979, pp. 207–230.
 26 https://doi.org/10.3102/10769986004003207.
- 27 25. Yen, W. M. Scaling Performance Assessments: Strategies for Managing Local Item
 28 Dependence. *Journal of Educational Measurement*, Vol. 30, No. 3, 1993, pp. 187–213.
 29 https://doi.org/10.1111/j.1745-3984.1993.tb00423.x.
- Stelmack, J., J. P. Szlyk, T. Stelmack, J. Babcock-Parziale, P. Demers-Turco, R. T.
 Williams, and R. W. Massof. Use of Rasch Person-Item Map in Exploratory Data Analysis:
 A Clinical Perspective. *The Journal of Rehabilitation Research and Development*, Vol. 41,
 No. 2, 2004, p. 233. https://doi.org/10.1682/JRRD.2004.02.0233.
- Bond, T. G., and C. M. Fox. Applying the Rasch Model: Fundamental Measurement in the
 Human Sciences. *Journal of Educational Measurement*, Vol. 40, No. 2, 2003, pp. 185–187.
 https://doi.org/10.1111/j.1745-3984.2003.tb01103.x.
- Miller, A., and S. Dishon. Health-Related Quality of Life in Multiple Sclerosis: The Impact
 of Disability, Gender and Employment Status. *Quality of Life Research*, Vol. 15, No. 2,
 2006, pp. 259–271. https://doi.org/10.1007/s11136-005-0891-6.
- 29. Duncan, P. W., R. K. Bode, S. Min Lai, and S. Perera. Rasch Analysis of a New StrokeSpecific Outcome Scale: The Stroke Impact Scale. *Archives of Physical Medicine and Rehabilitation*, Vol. 84, No. 7, 2003, pp. 950–963. https://doi.org/10.1016/S00039993(03)00035-2.
- Scheuneman, J. D., and C. A. Bleistein. A Consumer's Guide to Statistics for Identifying
 Differential Item Functioning. *Applied Measurement in Education*, Vol. 2, No. 3, 1989, pp.
 255–275. https://doi.org/10.1207/s15324818ame0203_6.

- Mantel, N., and W. Haenszel. Statistical Aspects of the Analysis of Data From
 Retrospective Studies of Disease. *JNCI: Journal of the National Cancer Institute*, Vol. 22,
 No. 4, 1959, pp. 719–748. https://doi.org/10.1093/jnci/22.4.719.
- Zwick, R., D. T. Thayer, and C. Lewis. An Empirical Bayes Approach to Mantel-Haenszel
 DIF Analysis. *Journal of Educational Measurement*, Vol. 36, No. 1, 1999, pp. 1–28.
 https://doi.org/10.1111/j.1745-3984.1999.tb00543.x.
- Martinková, P., A. Drabinová, Y.-L. Liaw, E. A. Sanders, J. L. McFarland, and R. M. Price.
 Checking Equity: Why Differential Item Functioning Analysis Should Be a Routine Part of
 Developing Conceptual Assessments. *CBE—Life Sciences Education*, Vol. 16, No. 2, 2017.
 https://doi.org/10.1187/cbe.16-10-0307.
- Schmitt, N., and D. M. Stuits. Factors Defined by Negatively Keyed Items: The Result of
 Careless Respondents? *Applied Psychological Measurement*, Vol. 9, No. 4, 1985, pp. 367–
 373. https://doi.org/10.1177/014662168500900405.
- 14 35. Campbell, D. T. Social Attitudes and Other Acquired Behavioral Dispositions. In
 15 Psychology: A study of a science. Study II. Empirical substructure and relations with other
 16 sciences. Volume 6. Investigations of man as socius: Their place in psychology and the
 17 social sciences., McGraw-Hill, New York, 2004, pp. 94–172.
- Pronello, C., and C. Camusso. Travellers' Profiles Definition Using Statistical Multivariate
 Analysis of Attitudinal Variables. *Journal of Transport Geography*, Vol. 19, No. 6, 2011,
 pp. 1294–1308. https://doi.org/10.1016/j.jtrangeo.2011.06.009.
- 37. Moraes, C., M. Carrigan, and I. Szmigin. The Coherence of Inconsistencies: Attitude–
 Behaviour Gaps and New Consumption Communities. *Journal of Marketing Management*,
 Vol. 28, No. 1–2, 2012, pp. 103–128. https://doi.org/10.1080/0267257X.2011.615482.
- 24 38. Sheeran, P. Intention-Behavior Relations: A Conceptual and Empirical Review. European 25 No. Review of Social Psychology, Vol. 12, 2002. 1. pp. 1 - 36.26 https://doi.org/10.1080/14792772143000003.
- Pronello, C., and J.-B. Gaborieau. Engaging in Pro-Environment Travel Behaviour
 Research from a Psycho-Social Perspective: A Review of Behavioural Variables and
 Theories. *Sustainability*, Vol. 10, No. 7, 2018, p. 2412. https://doi.org/10.3390/su10072412.
- 40. Hamidi, Z., and C. Zhao. Shaping Sustainable Travel Behaviour: Attitude, Skills, and
 Access All Matter. *Transportation Research Part D: Transport and Environment*, Vol. 88,
 2020, p. 102566. https://doi.org/10.1016/j.trd.2020.102566.
- 33 41. Bond, T. G., and C. M. Fox. Applying the Rasch Model. Psychology Press, 2007.
- Pronello, C., and P. Kumawat. Smartphone Applications Developed to Collect Mobility
 Data: A Review and SWOT Analysis. In *Advances in Intelligent Systems and Computing*,
 2021, pp. 449–467.
- 37