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

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

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

Cristina Pronello, corresponding author
Politecnico di Torino, Dipartimento Interateneo di Scienze, Progetto e Politiche del Territorio
Viale Mattioli, 39, 10125, Torino, Italy
Tel: +39-011-0905613; Fax: +39-011-0906450; Email: cristina.pronello@polito.it

Word count: 6,572 words text + 2 tables x 250 words (each) = 7,072 words

Submission Date: 26 July 2021

ABSTRACT

Ecological behaviour and the impact on environment are subjects of public concern and understanding individual behavioural measures to induce sustainable lifestyles is of extreme importance for policy makers to assess and promote sustainable mobility. To this end, a questionnaire with highly reliable items, evaluation of determinants and accurate measurement of ecological behaviour is a precondition for understanding the levers for a behavioural change. This paper aims at understanding whether the dichotomous Rasch model provides a legitimate measurement of General Ecological Behaviour (GEB) using a 26 items questionnaire as a valid tool to assess pro-environment behaviour of a large sample of users. A web questionnaire was 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 questionnaire is able to effectively measure pro-environment behaviour of travellers. Unidimensionality, perfect level of item reliability of 1, very high item separation of 34.22, 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 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 proposed approach allows also to measure the efficacy of the adopted policies.

Keywords: General Ecological Behaviour, Pro-environment Travel Behaviour, Dichotomous Rasch Model

INTRODUCTION

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

Therefore, various studies in literature used GEB to assess sustainable behaviour. Arnold 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 impact on humans and nature, the other represents automobile drivers' interests, such as proper road maintenance, allowing higher speed limits on freeways, and fighting gasoline-tax increases. Hergesell (*11*) examined differences in choosing the transport mode during the holidays through general level of environmental commitment across lifestyle domains and found that train users tend to be more environmentally committed compared to car users. Two versions of GEB questionnaire were proposed to assess pro-environment travel behaviour in an Italian region. A first version was proposed by Gaborieau and Pronello (*12*) based on Kaiser and Wilson (*10*), which we call GEB-40 (40 dichotomous items); the second version was proposed by (*13*) as an extended version of GEB-40, which we call GEB-51 (51 dichotomous items). One of the weaknesses of previous two Italian GEB versions (GEB-40 and GEB-51) was the inclusion of irrelevant and redundant items that were excluded in this study.

At best of our knowledge, the studies measuring GEB questionnaire using the Rasch model (*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 and Wilson (*10*) compared 686 Californian students and 445 Swiss participants; Gaborieau and Pronello (*12*) compare 131 Italian, 445 Swiss, and 247 Swedish participants; Hergesell (*16*) based on a sample of 349 German citizens, although the sample size is still within acceptable boundaries, 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 (*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:

- 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 pro-environment 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.

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

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 pro-environmental 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) model estimation and testing of GEB.

Survey Design

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

TABLE 1 Structure of GEB-26 Questionnaire

No.	Item description	Code
	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 crime	CS4
5	Sometimes I ride public transport without paying a fare	CS6(-)
	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.

Survey Administration and Sample Selection

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

Database Construction

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

The next step for constructing the data base was the check of missing values. Two 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, and they were considered as missing during the analysis to avoid any imputation, having a large database. The software Winsteps, used for the Rasch model, does not require complete data in order to make estimates, because it uses Joint Maximum Likelihood Estimation (JMLE) that is very flexible as regards estimable data structures. Waterbury (*18*) reports that Rasch model can 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 have complete data for corresponding records. Finally, the dataset is converted from polytomous scale to dichotomous scale by converting the first three categories, from 1 (completely disagree) to 3, to 1 “No”; and second three categories, from 4 to 6 (completely agree), to 2 “Yes”.

Rasch Model as a Measure of General Ecological Behaviour

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

$$B_n - D_i = \ln(P_{ni}/P_{ni}) \quad (1)$$

where

B_n = ability of a specific person n;

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

D_i = difficulty of a specific item i;
P_{ni} = probability of person n correctly answering item i; and
ln = "log-odds units" (logits), which is a natural logarithm.

The DRM specifies the probability, P, that the person *n* with ability B_n succeeds on item *i* of difficulty D_i.

The key Rasch model requirements are unidimensionality, local independence, persons-invariant item estimates/person parameter separability, and item-invariant person estimates/item parameter separability.

For the parameter estimation for DRM, Winsteps Rasch Analysis program was used. Winsteps implements two methods of estimating Rasch parameters from ordered qualitative observations: JMLE also known as UCON (Unconditional Maximum Likelihood Estimation) (*20*) and PROX (Normal Approximation Algorithm) devised by Cohen (*21*).

Rasch Measures and Model fit

Rasch model fits are used to examine the unidimensionality of the latent trait to measure attitude towards GEB. Unidimentionality is evaluated using: 1) point-biserial correlation 2) fit statistics, 3) Principal Component Analysis of Residuals, and 4) local independence.

Point-biserial Correlation Point-biserial correlation is a useful diagnostic indicator of data miscoding or item mis-keying: negative or zero values indicate items or persons with response strings that contradict the variable. Li et al. (*22*) suggests that point-measure correlation larger than .3 indicate that items are measuring the same construct.

Fit Statistics Rasch model provides two indicators of misfit: INFIT and OUTFIT. Since the ZSTD value is based on the MNSQ, as reported by Boone et al. (*23*), we first examine the MNSQ for evaluating fit. If the MNSQ value lies within an acceptable range, we ignore the ZSTD value. According to Boone et al. (*23*), INFIT and OUTFIT mean-square fit statistics between 0.5-1.5 represents productive items. For mathematical formulation of point-biserial correlation, INFIT, OUTFIT, and ZSTD refer to (*12*).

Principle Component Analysis of Residuals (PCAR) Unidimensionality was checked through PCAR. According to Reckase (*24*) unidimensionality is hold if: a) the amount of variance explained by measures is > 20%; b) unexplained variance of the eigenvalue for the first contrast is < 3; and unexplained variance accounted by first contrast is < 5%.

Local Independence Local independence means that after the contribution of the latent trait(s) to the data is removed, all that is left is random noise (*25*). A correlation of r=0.40 among items is low dependency.

Besides these, Rasch model assumptions include assessing *reliability and separation* of measures, *differential item functioning*, evaluation of item difficulty using *Write map* to evaluate construct validity.

Reliability It ranges from 0 to 1 and the higher is better (*26*). Bond and Fox (*27*) suggested value between 0.6-0.8 is acceptable.

Separation index Seperation 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*).

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

RESULTS

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

Point-biserial Correlations

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

- *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).

Fit statistics

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

Principle Component Analysis of Residuals

First, the amount of variance explained by measures is 34.2% (11.5% of raw variance explained by persons and 22.7% of raw variance explained by items) which is larger than the requirement of 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.

The loading of items on the first contrast of the residual based PCA are shown in Figure 1, showing that this possible sub-dimension is formed by two items, A (AE6_REVC), and B (CS6_REVC). Items A and B have the largest loadings, quite far away from the general cluster created by the other items, and the eigenvalue of first contrast is 2.14 (~2 items). To see the items corresponding to the letter of alphabet represented in Figure 1, refer to Table 2.

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

Entry No.	Total Score	Measure	Model S.E.	Infit		Outfit		Point-bis. Corr.		Exact Match (%)		Item
				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	H .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	l .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	-

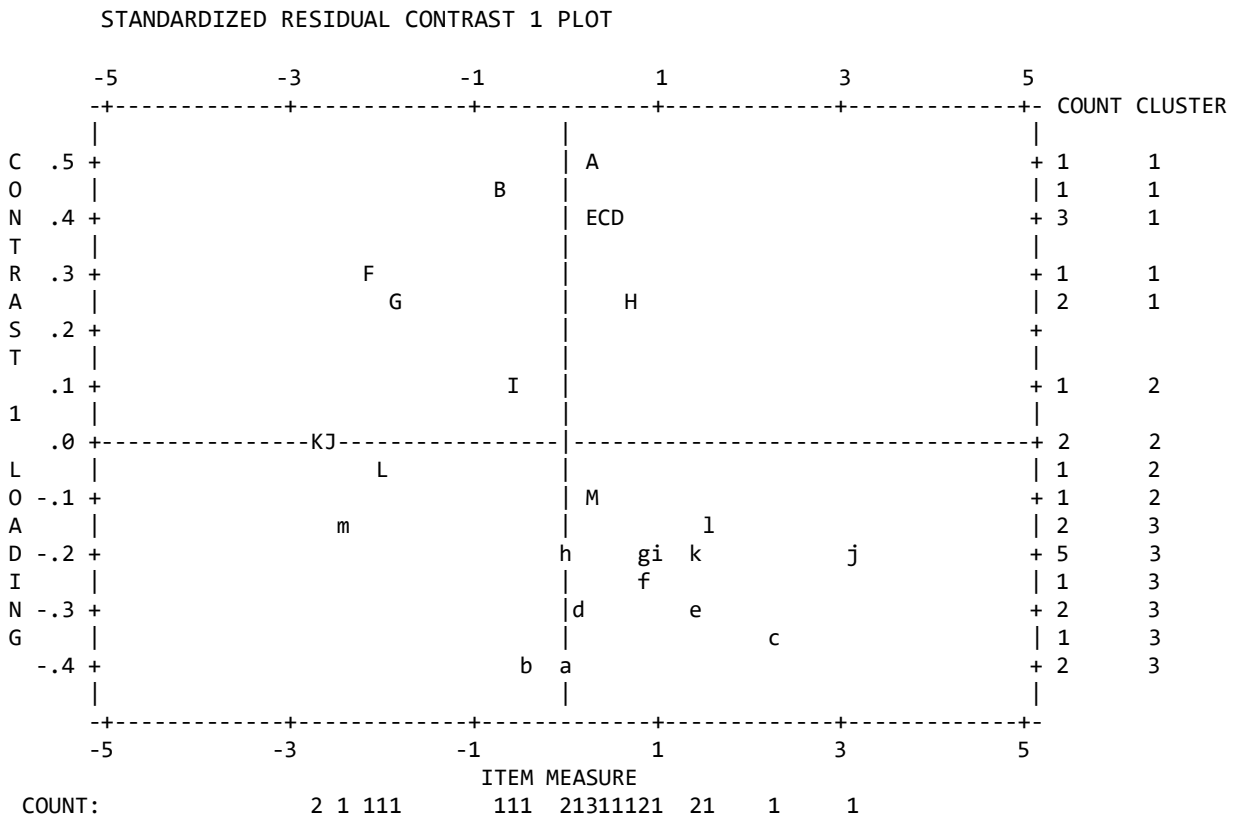

FIGURE 1 Item loadings on the first contrast

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

Local independence

According to the Linacre guidelines[2] all items correlation is <0.4, hence no item residuals are correlated, respecting the local independence assumptions of Rasch analysis.

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*).

Item separation is very high, equal to 34.22, and represent excellent level of separation (*29*). With this large person sample, the item difficulties are estimated exceedingly precisely and validating the GEB construct validity (>3).

[2] https://www.winsteps.com/winman/table23_99.htm.

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.

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

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:

- 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.

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.

Unidimensionality has been evaluated utilizing Rasch fit statistics, as well as PCAR and point-biserial correlations. Notably, all these tests of the measure's dimensionality suggest the items lie on one trait as hypothesized during survey design. Therefore, it can be recommended to 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 is not threatening the validity of the scale, so that it is not suggested to delete it. The fact that item AE6_REVC was the only item with poor fit warrants further investigations as it offers potential insights into the structure of GEB. It is well known that negatively coded items, especially if there are only a few and located at the end of the questionnaire, may be confusing for the respondents (*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 rather a (partly) different construct in its own right. Moreover, local independence, reliability, and separation indexes assumptions were confirmed with good Rasch measures validity.

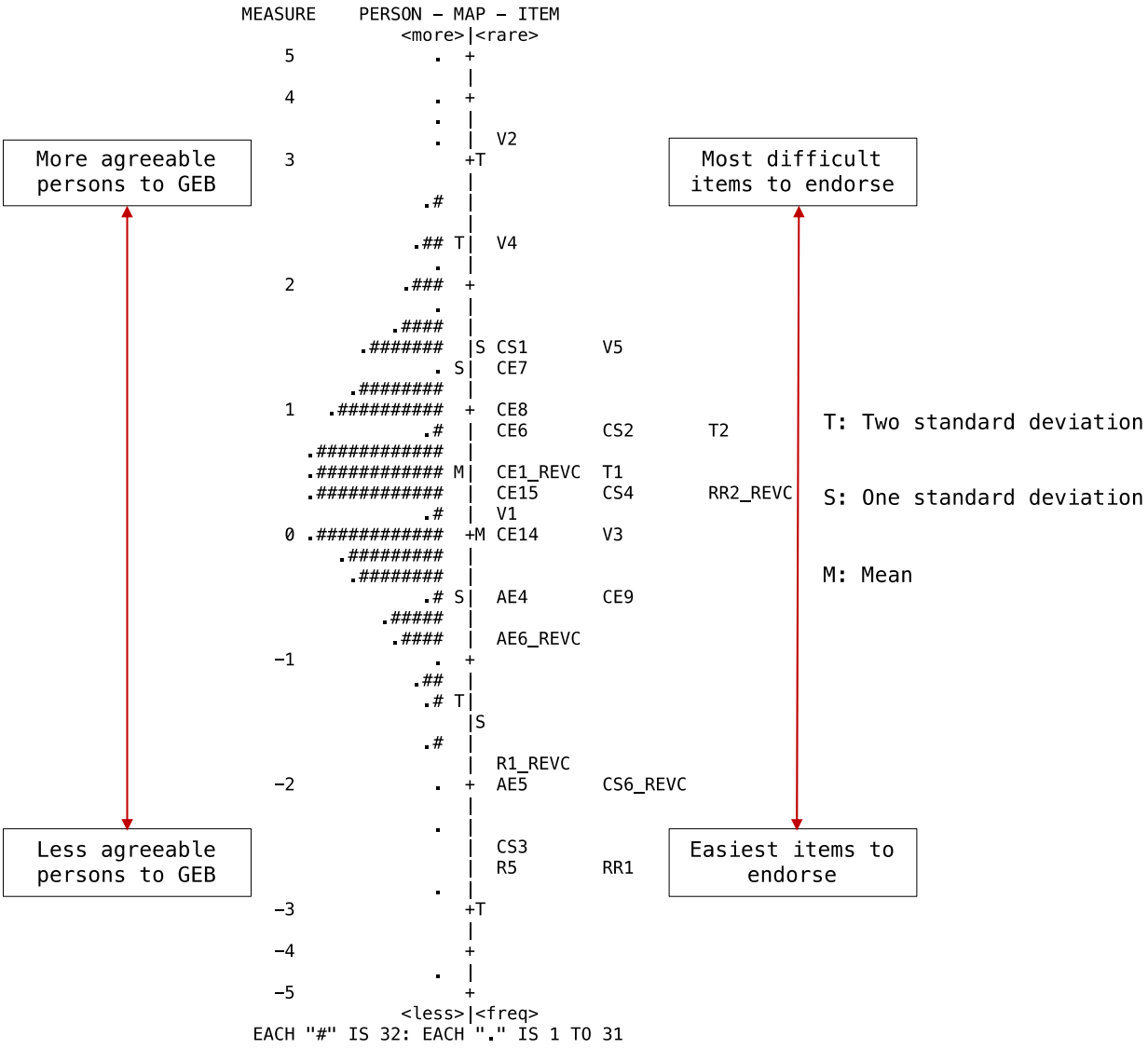

FIGURE 2 Write map

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

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

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

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 people's daily lifestyles (*2*). A frequent question arising is: do what people report to protect the environment converge with their environmental impact? If, theoretically speaking, this could hold 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 (32.48%) use trip chain followed by 1156 (27.45%) using car, 729 (17.31%) using public transport, 330 (7.83%) walking, and 310 (7.36%) cycling. Looking into trip chain, car as driver is used by the highest percentage of respondents, 1333 (31.65%), followed by 1096 people (26.02%) using 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 may reflect actual environmental impact in some other contexts such as in electricity consumption, as reported by Arnold et al. (*9*), but they do not apply in transport sector by looking at results and as shown in previous studies (*36*). This is referred as attitude-behaviour gap (*37*) or behaviour 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 correspondent to those users who use soft modes (walking or bike) for their most frequent trip, followed by public transport (regional train, bus, tram or metro) and, then, private motorised vehicles (car or motorbike). One reason of this contradiction might be that the trip chain was excluded by Gaborieau and Pronello (*12*) and the sample was smaller (108 users). This discrepancy will be further investigated in the continuation of the research.

It should also be recalled that the sample sizes in previous studies – in Italian context (GEB-40 and GEB-51), in Swedish and Swiss context (*15*), and in Californian context (*10*) – were comparatively too small, although still within acceptable boundaries, according to Linacre (*17*). 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 previous studies was formed thanks to a stratified sampling plan. Thus, different results may be observed when the sample follows the snowball sampling approach and the participants are, as in this case, younger and/or with a bit lower educational level. Finally, it needs to be emphasized that even excellent internal validity is no assurance that a given scale will also exert good external 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*).

Further research is needed to deepen our understanding of the GEB and to devise appropriate measurement instruments. There was no evidence that individuals with diverging 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 be answered by respondents with high and low ability. Overlapping items measure different elements with different levels of difficulty (*41*), hence we do not suggest to exclude items to design a new survey by looking only at redundancy of items in write map. Some recommendations are worthy to be given for improving the scale. Firstly, more items could be selected with high or low 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. This is important because limited differentiation capabilities may attenuate existing effects of measuring ecological behaviour. The GEB-26 might not be capable of detecting strong effects potentially attributable to interventions based on ecological behaviour in terms of larger person ability range due to weakness of questionnaire design; in fact, we obtained person measure reliability equal to .67 and person separation equal to 1.44, which is acceptable but not excellent. Hence, GEB researchers would profit from more sensitive measurement instruments capable of detecting differences between individuals who are high and low in ecological behaviour. Furthermore, we do not suggest excluding any item by looking only at the dichotomous scale measurement. The item exclusion will be further decided after measuring the original 6 scale 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 suggested by Linacre[3], the scale with more categories is expected to give better and higher person 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 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.

The final aim is to propose proper public targeted policies to induce people to sustainable travel choices. For that, a well-planned and environment-focused transport education policy can play a role to educate people and make them aware about their environmental footprints and motivate them to behave ecologically and sustainably. In this regard, the transport policies together 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 can also play a role for giving incentives to promote sustainable mobility or engaging them in pro-environment ecological behaviour with the help of smartphone apps, as these days the apps are 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.

good, validated, and reliable measurement of ecological behaviour would let public bodies to measure the efficacy of adopted policies.

AUTHOR CONTRIBUTION STATEMENT

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

REFERENCES

1. 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.
2. 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.
3. 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. 149–159. https://doi.org/10.1016/S0272-4944(05)80147-1.
4. 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.
5. C Stern, P. Toward a Coherent Theory of Environmentally Significant Behaviour. *Journal of Social Issues*, Vol. 56, No. 3, 2000, pp. 407–424.
6. 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.
7. 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.
8. 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.
9. 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.
10. Kaiser, F. G., and M. Wilson. Assessing People's General Ecological Behavior: A Cross-Cultural 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.
11. 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.
12. 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.
13. Duboz, A. *The Intention to Use Real-Time Multimodal Information to Change Travel Behaviour The Use of Psychosocial Variables for the Market*. Politecnico di Torino, Italy, 2018.
14. 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.
15. 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.
16. Hergesell, A. Exploring the Relationship between General Environmental Behaviour and

Holiday Transport Mode Choice. *BEST EN Think Tank XIV: Politics, Politics, Policy and Governance in Sustainable Tourism*, 2014, pp. 142-149.

17. 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.
19. Pallant, J. F., and A. Tennant. An Introduction to the Rasch Measurement Model: An Example Using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*, Vol. 46, No. 1, 2007, pp. 1–18. https://doi.org/10.1348/014466506X96931.
20. 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.
21. 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.
22. 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.
23 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. Reckase, M. D. Unifactor Latent Trait Models Applied to Multifactor Tests: Results and Implications. *Journal of Educational Statistics*, Vol. 4, No. 3, 1979, pp. 207–230. https://doi.org/10.3102/10769986004003207.
25. Yen, W. M. Scaling Performance Assessments: Strategies for Managing Local Item Dependence. *Journal of Educational Measurement*, Vol. 30, No. 3, 1993, pp. 187–213. https://doi.org/10.1111/j.1745-3984.1993.tb00423.x.
26. 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.
27. 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.
28. 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 Stroke-Specific 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/S0003-9993(03)00035-2.
30. 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.

31. 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.
32. 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.
33. 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.
34. 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.
35. Campbell, D. T. Social Attitudes and Other Acquired Behavioral Dispositions. In *Psychology: A study of a science. Study II. Empirical substructure and relations with other sciences. Volume 6. Investigations of man as socius: Their place in psychology and the social sciences.*, McGraw-Hill, New York, 2004, pp. 94–172.
36. 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.
38. Sheeran, P. Intention—Behavior Relations: A Conceptual and Empirical Review. *European Review of Social Psychology*, Vol. 12, No. 1, 2002, pp. 1–36. https://doi.org/10.1080/14792772143000003.
39. 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.
41. Bond, T. G., and C. M. Fox. *Applying the Rasch Model*. Psychology Press, 2007.
42. 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.