Development and validation of a photograph-based instrument to assess nutrition literacy: the NUTLY project

Sofia Sousa, Gabriela Albuquerque, Milton Severo, Ana Rute Costa, Pedro Moreira, Nuno Lunet, Patrícia Padrão

PII: S0195-6663(24)00178-8

DOI: https://doi.org/10.1016/j.appet.2024.107377

Reference: APPET 107377

To appear in: *Appetite*

Received Date: 5 October 2023

Revised Date: 7 April 2024

Accepted Date: 24 April 2024

Please cite this article as: Sousa S., Albuquerque G., Severo M., Costa A.R., Moreira P., Lunet N. & Padrão P., Development and validation of a photograph-based instrument to assess nutrition literacy: the NUTLY project, *Appetite*, https://doi.org/10.1016/j.appet.2024.107377.

Development and validation of a photograph-based instrument to assess nutrition literacy: the NUTLY project

Sofia Sousa[1,2], Gabriela Albuquerque[1,2], Milton Severo[1,2,3], Ana Rute Costa[1,2,4], Pedro Moreira[1,2,5,6], Nuno Lunet[1,2,4 *], Patrícia Padrão[1,2,5 *]

[1] EPIUnit – Instituto de Saúde Pública da Universidade do Porto

[2] Laboratório para a Investigação Integrativa e Translacional em Saúde Populacional (ITR)

[3] Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto

[4] Departamento de Ciências da Saúde Pública e Forenses, e Educação Médica, Faculdade de Medicina da Universidade do Porto

[5] Faculdade de Ciências da Nutrição e Alimentação da Universidade do Porto

[6] Centro de Investigação em Atividade Física, Saúde e Lazer da Universidade do Porto

* Contributed equally to the manuscript

Corresponding author:

Patrícia Padrão

Faculdade de Ciências da Nutrição e Alimentação da Universidade do Porto

Rua do Campo Alegre, 823 / 4150-180 Porto, Portugal

Telephone: +351 225074320 | Fax: +351 225074329

Email: patriciapadrao@fcna.up.pt

Abstract

Most instruments measuring nutrition literacy evaluate theoretical knowledge, not necessarily reflecting skills relevant to food choices. We aimed to develop and validate a photograph-based instrument to assess nutrition literacy (NUTLY) among adults in Portugal. NUTLY assesses the ability to distinguish foods with different nutritional profiles; from each of several combinations of three photographs (two foods with similar contents and one with higher content) participants are asked to identify the food with the highest energy/sodium content. The NUTLY version with 79 combinations, obtained after experts/lay people evaluations, was applied to a sample representing different age, gender and education groups (n=329). Dimensionality was evaluated through latent trait models. Combinations with negative or with positive small factor loadings were excluded after critical assessment. Internal consistency was measured using Cronbach's alpha and construct validity by comparing NUTLY scores with those obtained in the Medical Term Recognition Test and the Newest Vital Sign (NVS), and across education and training in nutrition/health groups. The cut-off to distinguish adequate/inadequate nutrition literacy was defined through ROC analysis using the Youden index criterion, after performing a Latent class analysis which identified a two-class model to have the best goodness of fit. Test-retest reliability was assessed after one month (n=158). The final NUTLY scale was unidimensional and included 48 combinations (energy: 33; sodium: 15; α=0.74). Mean scores (±standard deviation) were highest among nutritionists (39.9±4.4), followed by health professionals (38.5±4.1) and declined with decreasing education ($p<0.001$). Those with adequate nutrition literacy according to NVS showed higher NUTLY scores (37.9±4.3 vs. 33.9 ±6.9, $p<0.001$). Adequate

nutrition literacy was defined as a NUTLY score≥35 (sensitivity: 89.3%; specificity: 93.7%). Test-retest reliability was high (ICC=0.77). NUTLY is a valid and reliable nutrition literacy measurement tool.

Keywords: nutrition literacy; NUTLY; Portugal; validation study.

1. Introduction

Health literacy can be defined as "the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions" (Ratzan & Parker, 2000). Health literacy is a stronger predictor of health than age, income, employment status, education level or ethnicity (World Health Organization, 2013), and it was shown to be lower in Portugal than in other European countries (Paiva, et al., 2017).

Nutrition literacy is a key component of health literacy, and has been recognized as an important determinant of food choice (Chen & Antonelli, 2020). Some studies suggest that higher nutrition literacy may be associated with healthier dietary practices (Elmskini, et al., 2024; Spronk, Kullen, Burdon, & O'Connor, 2014; Taylor, Sullivan, Ellerbeck, Gajewski, & Gibbs, 2019). However, they also indicate that increasing factual knowledge alone may not be sufficient to improve dietary behaviour (Jezewska-Zychowicz & Plichta, 2022; Koch, Hoffmann, & Claupein, 2021). Nutrition literacy, which comprises both knowledge and skills, can influence consumers' ability to identify healthy foods (Grunert, Wills, & Fernández-Celemín, 2010), being considered as essential in the prevention of diet-related noncommunicable diseases (NCDs) (World Health Organization, 2022). The term "nutrition literacy" has been defined as "the degree to which individuals are able to obtain, process and understand nutrition and diet information, as well as access services needed to make adequate nutrition decisions". This definition clearly focuses on the set of knowledge and skills with practical relevance for healthy food choices (Vettori, Lorini, Milani, & Bonaccorsi, 2019). However, most instruments assess knowledge about theoretical concepts in

nutrition, which may not translate into practical knowledge or the capacity to understand and apply nutrition information at its different levels, in order to make appropriate dietary decisions that promote health (Carbone & Gibbs, 2013; Spronk, et al., 2014). Furthermore, the available instruments strongly rely on the participants' ability to read, write or perform calculations, limiting their use in population groups with the lowest educational levels, especially exposed to poor health outcomes (Zajacova & Lawrence, 2018).

There are a few nutrition literacy measuring tools already validated for the Portuguese population. The General Nutrition Knowledge Questionnaire (GNKQ) (Almeida-Souza, 2009; Ferro-Lebres, Moreira, & Ribeiro, 2014) addresses theoretical knowledge about nutritional facts and recommendations and the Nutrition Literacy Assessment Instrument (NLAI) evaluates knowledge on nutrients, portions of food, groups of the Portuguese Food Wheel and food labels (Azevedo, et al., 2019). A Portuguese version of the Newest Vital Sign (NVS) was also validated (Martins & Andrade, 2014). Although being described as a health literacy screening tool, it focuses on reading and interpretation of a nutrition label, therefore fitting into the nutrition literacy concept (Vettori, et al., 2019). Like other instruments, these present an essentially theoretical perspective, not assessing the ability to apply nutrition knowledge in the selection of foods. Also, the complexity of the questions may cause them to be of limited application in individuals with low levels of general literacy and/or numeracy. Therefore, the aim of this study was to develop and validate a photograph-based instrument for measuring nutrition literacy (NUTLY) in adults residing in Portugal. The instrument aims to address nutrition-related knowledge and skills with practical relevance for food choice, by exploring the ability to distinguish foods with different nutritional

profiles, through an approach similar to food decisions from daily life. Furthermore, by relying on photographs rather than words or numbers, NUTLY may facilitate the assessment of individuals across a broader range of educational levels.

2. Methods

2.1. Rationale for the NUTLY instrument

The rationale behind the NUTLY instrument took into consideration two main aspects: (1) which nutritional parameters to evaluate; and (2) how to evaluate the ability to distinguish foods based on a given nutritional parameter, taking advantage of the use of photographs of foods.

Regarding the first aspect, four nutritional parameters were considered, with the aim of addressing the nutrients with strong evidence on the relationship with diet-related NCDs: (1) energy; (2) sodium; (3) saturated- and trans-fat; and (4) sugars. However, we decided to exclude saturated- and trans-fat and sugars, due to the importance of these nutrients to the total energy value, which would add a correlation between these variables, thus constituting a confounding factor in the evaluation of these parameters. As such, it was decided that the NUTLY tool would target energy and sodium, whose excessive intakes are very frequent and strongly linked, respectively, with overweight/obesity and hypertension, which are two key risk factors for NCDs, at both national and global levels (World Health Organization, 2018).

Regarding the second aspect, this tool was idealized to measure nutrition literacy, which entails competencies related to the acquisition, comprehension and application of knowledge needed to make adequate nutrition decisions. As such, we designed this instrument with the particularity of mimicking as much as possible the real context of food choices, so that it could reflect the dimensions behind this construct. For this, the NUTLY

instrument was thought to consist of the consecutive presentation of sets of three photographs of different food items (hereafter referred to as "combinations"), comprising two food items with similar estimated contents and one with higher content, considering one specific nutritional parameter. For each combination of three photographs of foods, the participants would be asked to identify the food item with the highest amount of either "energy (calories)" or "salt", depending on the parameter being evaluated in each combination. Two examples of such combinations are presented in **Supplementary Figure 1**.

2.2. Development of the NUTLY instrument

The development of the NUTLY instrument comprised four main steps: (1) selection of the food items to be photographed; (2) production, evaluation and continuous improvement of the photographs; (3) preparation of a nutritional composition database for the selected food items; and (4) assembling of the first version of the NUTLY instrument. A flowchart summarizing the development of the NUTLY instrument is presented in **Figure 1a**.

2.2.1. Step 1: Selection of the Food Items to be photographed

A total of 90 foods or culinary preparations were selected taking into account the following criteria: (1) culturally familiar to people living in Portugal; (2) usually consumed in different meals throughout the day or food contexts; (3) easily recognizable; and (4) diverse in terms of composition. These food items were categorized into 11 food groups: (1) main dishes; (2) soups; (3) sandwiches; (4) savoury pastries and snacks; (5) sweet cakes and pastries; (6)

desserts; (7) confectionery; (8) cookies and biscuits; (9) bread; (10) charcuterie and cheese; and (11) fruit. Serving sizes to be photographed were defined using Portuguese reference portions(Goios, Martins, Oliveira, Afonso, & Amaral, 2016; PortFIR – Grupo de Trabalho Porções, 2014).

2.2.2. Step 2: Production, Evaluation and Continuous Improvement of the Photographs

All photographs were produced using a CANON® EOS 2000D photographic camera with an EF-S 18-55 mm f/3.5-5.6 DC lens. A tripod was used to provide stability and a ring flash was added to the lens to eliminate shadows. All food items were placed at the centre of a white plate. A 15 cm ruler and a pen were also added at the top and right of the image, respectively, to normalize the perception of size. As necessary, image corrections (e.g., saturation, contrast, brightness) were made, in order to improve and/or standardize the visual aspect of the photographs.

In a first round, from the 90 food items initially selected, a total of 55 food items from all food groups were photographed using three different angles (10°, 45° and 90° from the horizontal plane) and two background colours (blue and grey). These photographs were presented to a panel of nine experts in Nutrition and Public Health, who performed the *phase 1 of external evaluation*. In this phase, experts were asked to identify each food item, and to compare the different angles and background colours in terms of food item recognition (i.e., at what level did the subject recognize the food(s) in the photograph, with a score ranging between 1 “Does not recognize, recognizes with extreme difficulty or considers that the photograph is extremely ambiguous” and 5 “Recognizes easily, there is

no ambiguity, and it does not seem needed to improve the quality of the photograph or of the food(s) presented") and photo aesthetics (by selecting the preferred photograph) (**Supplementary Figure 2a**). Exceptionally, the soups were only presented in the 45° and 90° angles, because the 10° angle made their content not visible.

Based on the feedback collected from this panel, the best scored angle and background colour were selected, and in the second round of photographs production, the remaining 35 food items were photographed, using only the selected angle and background colour. All photographs (including those that had already been evaluated in the first phase of external evaluation) were assessed in the *phase 2 of external evaluation*. In this phase, two panels were included: (1) an expert panel, composed of 12 experts in Nutrition and Public Health; and (2) a lay panel, constituted by 10 participants, including men and women and representing different age groups and geographic locations in Portugal. Both panels were asked to identify and describe each item in detail, classify the perceived ease of recognition (*i.e.*, how easy it was to recognize the food item, on a scale from 1 to 5), quantify the portions presented (by estimating, in grams, the quantity of the food item(s) presented) and classify, according to their individual perception, if the portion presented in the photograph was small, medium or large (to test whether each photograph could give an increased or decreased perception of the real portion size). In the case of multi-component main dishes (*e.g.*, steak with rice, french fries and salad), panel members were asked to quantify each component individually (**Supplementary Figure 2b**). Both phases of external evaluation were conducted using a self-administered questionnaire.

184 At the end of each of these phases of external evaluation, photographs were improved
185 following a set of decision rules (**Supplementary Figure 3**).
186
187 *2.2.3. Step 3: Preparation of a Nutritional Composition Database for the selected food items*
188 A nutritional composition database was produced, containing the energy and sodium
189 contents per reference portion of each one of the photographed food items, being last
190 updated on 21 April 2022 (**Supplementary Table 1**). Nutritional data were collected from
191 the Portuguese Food Composition Table, coordinated by the Portuguese National Institute
192 of Health *Doutor Ricardo Jorge*, under the authority of the Portuguese Ministry of Health
193 (Instituto Nacional de Saúde Doutor Ricardo Jorge, 2021). This database compiles data from
194 different sources in full compliance with EuroFIR's requirements. In the case of composite
195 dishes, nutritional contents are calculated using the EuroFIR method for recipe calculation
196 (EuroFIR, 2008). The foods included in this database (including the composed dishes) take
197 into account consumption data from the most recent National Food, Nutrition and Physical
198 Activity Survey, the IAN-AF 2015-2016 (Lopes C, et al., 2017). In the case of industrially-
199 produced foods with no data available in the Portuguese Food Composition Table, the
200 nutritional information from food labels of market-leading brands.
201
202 *2.2.4. Step 4: Assembling of the First Version of the NUTLY instrument*
203 For the first version of the NUTLY, 77 combinations of three photographs (55 for the energy
204 parameter and 22 for the sodium parameter) were created using two criteria: (1) in each
205 combination, the food items to be compared are commonly consumed in the same meal or

food context, regardless of whether or not they belong to the same food group; and (2) the combinations are diverse in terms of the difficulty of the comparison, calculated for this purpose as the percentual difference between the option with the highest content (i.e., the correct answer) and the other two options with lower contents. The lower this difference was (i.e., the closer the contents of the three foods were), the higher the difficulty of the set. Such differences varied between 19% and 83% (for the energy parameter), and between 35% and 100% (for the sodium parameter) (**Supplementary Figure 4**).

In order to address potential variations in the composition of the culinary preparations, we added a note highlighting that "The foods presented in the photographs correspond to the usual recipes and culinary practices in the general Portuguese population, including the ingredients commonly added in the average consumed amounts.", placed in the beginning of the instrument (i.e., immediately before the energy parameter) and repeated in the middle of the instrument (i.e., immediately before the sodium parameter). In the specific case of the sodium combinations, we excluded main dishes due to the potential variations in the quantity of added salt during cooking or at the table (discretionary salt), and soups were only used in two combinations where the differences in sodium contents between options clearly outweighed these potential variations in the salt addition.

2.3. Validation Study

The validation study aimed to assess validity and reliability measures of the NUTLY and comprised four main steps: (1) external evaluation of the first version of the NUTLY instrument, to assess content validity; (2) pre-test, to evaluate its comprehension (the

extent to which the respondents interpret and answer the questions in alignment with the intended purpose of the instrument); (3) application of the NUTLY questionnaire, to measure internal consistency reliability, construct validity, and responsiveness (sensitivity and specificity); and (4) follow-up for assessment of test-retest reliability. A flowchart summarizing the validation of the NUTLY instrument is presented in **Figure 1b**.

2.3.1. Step 1: External Evaluation of the First Version of the NUTLY Instrument – Content Validity

The combinations created at the end of the development stage underwent a new external evaluation by two panels (experts in Nutrition and Public Health and lay people) (**Supplementary Figure 5**). In addition to selecting the photograph corresponding to the food item that they considered to have the highest amount of energy/sodium, both panels were also asked to classify each comparison according to their perception of how difficult it was to identify the item with the highest content. Furthermore, the experts were asked to evaluate the relevance of each combination (on a scale of 1 to 4), considering the measurement aim of the instrument. Commentaries/suggestions were also collected from both panels. For each combination, the placement of the photographs in the screen (left, centre, right), as well as the order in which the combinations appeared throughout the instrument, were randomly defined.

The frequency of correct answers was compared between panels, and associated with the perceived level of difficulty, in order to understand if the instrument was performing as expected (*i.e.*, experts scoring higher than lay people, and frequency of correct answers

250 decreasing with increasing levels of difficulty). The relevance score was used to calculate
251 the Content Validity Index (CVI) of each combination, as follows: CVI = number of
252 classifications “3” or “4”/total number of experts. Combinations with a CVI ≥ 0.75 were
253 considered to have excellent content validity (Polit, Beck, & Owen, 2007). Considering the
254 feedback from these panels, two new combinations were added to the set previously
255 defined to evaluate the sodium parameter.
256
257 *2.3.2. Step 2: Pre-test of the NUTLY Questionnaire – Comprehension*
258 The NUTLY instrument comprising 79 combinations (55 for energy and 24 for sodium) was
259 prepared as an online questionnaire to be self-administered. All the instrument’s
260 instructions were presented in written and audio formats.
261 The questionnaire (containing the NUTLY instrument, sociodemographic and health/diet-
262 related questions and two other instruments to assess nutrition and health literacy) was
263 then pre-tested in a sample of nine adults of different age groups, gender and education
264 levels, in order to evaluate the instrument’s comprehension, as well as to identify potential
265 lack of clarity of the provided instructions and questions.
266
267 *2.3.3. Step 3: Application of the NUTLY Questionnaire – Internal Consistency, Construct*
268 *Validity, Sensitivity and Specificity*
269 2.3.3.1. Sampling and data collection
270 Eligible participants were adults (aged 18 years or older) without visual or cognitive
271 impairments that could preclude the correct application of the instrument. Three different

sampling strategies were adopted, in order to cover different age, gender and education level groups: (1) the questionnaire was sent by email to the academic community of the University of Porto, which included students, teaching and non-teaching staff; (2) computers were made available in the premises of the faculties which agreed to collaborate, that way reaching potential participants that did not use the institutional email regularly, that did not have easy access to a digital device with internet, or that were not independent users; and (3) computers were made available in a recreational senior centre in the North of Portugal, where all users and staff members were invited to participate. The questionnaire was self-administered and could be fulfilled using different devices, namely computer, tablet, or smartphone. For the participants to answer the questionnaire in computers provided by the research team, assistance to solve any technical problem was provided upon request. The sample size estimation was based on the quality criteria for measurement properties of health status questionnaires proposed by Terwee et al. These authors propose a "rule of thumb approach" by multiplying the number of variables in the questionnaire per seven subjects, with a minimum number of 100 subjects, to ensure the stability of variance-covariance matrix (Terwee, et al., 2007). As such, a sample size of 315 was estimated, based on a predicted final instrument of approximately 45 combinations.

Data collection occurred between October 2022 and January 2023. Before the questions comprising the NUTLY instrument, participants were asked which type of device (computer, tablet, smartphone or others) was being used to answer the questionnaire. Sociodemographic (gender, age, education level, higher education in health and/or nutrition) and health/diet-related information [self-reported current weight and height

(which were then used to calculate body mass index), previous diagnosis of diet-related NCDs, responsibility for buying or cooking food at home, importance given to the nutritional composition of the foods consumed, history of consultation with a nutritionist] were also assessed. After the NUTLY instrument, the final questions corresponded to two instruments validated for the Portuguese population that were applied to assess health literacy: the Medical Term Recognition Test (METER) (Paiva, et al., 2014) and the NVS (Martins & Andrade, 2014), the latter being composed of questions specifically related to the area of nutrition.

2.3.3.2. Internal consistency

The NUTLY instrument was composed of two main groups of items (i.e., combinations of three photographs), each one corresponding to the two nutritional parameters to evaluate: energy and salt. As such, latent trait models were used to assess the dimensionality of the scale, to test whether it was organized into one dimension (i.e., unidimensional scale) or two dimensions (i.e., energy dimension and salt dimension) and also to identify which items (i.e., combinations) belonged to each dimension, if applicable. Item reduction was made at this stage, through the exclusion of combinations with a small standardized factor loading (lower than 0.2), combined with critical assessment by the research team, in order to obtain a final version with a high level of internal consistency (Cronbach's alpha between 0.70 and 0.95 (Terwee, et al., 2007)) ~~**(Supplementary Table 2)**~~.

2.3.3.3. Construct validity

316 Construct validity was tested using two complementary criteria of convergent validity: (1)
317 comparison with two health and nutrition literacy instruments validated for the Portuguese
318 population, the METER; and NVS, respectively; and (2) association with educational level
319 and academic training in nutrition and health, under the assumption that nutrition literacy
320 is higher among nutritionists, followed by those with health academic training and gradually
321 decrease with lowering education levels, based on previous literature (Jansen, et al., 2018;
322 Mohsen, Sacre, Hanna-Wakim, & Hoteit, 2022; van der Heide, et al., 2016).
323
324 2.3.3.4. Sensitivity and specificity
325 Under certain conditions, it is possible to estimate the sensitivity and specificity even
326 though no gold standard is available (Walter & Irwig, 1988). Under the assumption that each
327 NUTLY item is an imperfect test (sensitivity and specificity below 100%), a two-class latent
328 class analysis model was confirmed to have the best goodness of fit, *i.e.,* the lowest Bayesian
329 Information Criterion (BIC), and was used to estimate the sensitivity and specificity of each
330 item, and the *a posteriori* classification of the outcome of each individual. The latter was
331 then used to calculate the area under the receiver operating characteristic (ROC) curve for
332 the total score (defined as the sum of correct answers). The Youden index criterion, i.e., the
333 score in which the Youden's J statistic (sensitivity + specificity - 1) was maximized, was used
334 to identify a threshold (cut-off value) to distinguish between “adequate” and “inadequate”
335 nutrition literacy that maximised both sensitivity and specificity.
336
337 *2.3.4. Step 4: Follow-up Evaluation – Test-rest Reliability*

The participants who authorised a second contact via e-mail and/or SMS were evaluated again with the NUTLY instrument, between one and two months after the first evaluation. The instrument's test-retest reliability was determined using the intra-class correlation coefficient (ICC) fixed raters single measure.

2.4. Statistical analysis

Frequencies and descriptive statistics were used to describe the participants' characteristics, as well as the parameters assessed during the external evaluations of the photographs. T-test and one-way ANOVA were used to test for differences in mean NUTLY scores, according to the participants' characteristics. Spearman's correlations were calculated between the NUTLY scores and METER and NVS. Specific statistical analyses performed in each step of the validation process were described in the respective section. The R software version 4.3.0 and Stata version 15.1 (StataCorp LP, College Station, TX) were used.

3. Results

3.1. Development of the NUTLY instrument

3.1.1. Phase 1 of External Evaluation of the Photographs

From the 55 food items photographed in the first round of photographs production, almost all (96.4%) were correctly identified by all panel members. The oblique angle (45° from the horizontal plane) obtained a median score of 5.0 (IQR = 0.0) on the perceived ease of recognition scale in 47 of the 55 food items evaluated (85.5%), whereas this frequency was lower for the 90° and 10° angles: 67.3% and 17.3%, respectively. The 45° angle was also considered the most aesthetically pleasing in the majority (81.8%) of the items evaluated. Regarding the background colour, the grey option was preferred in 92.7% of the items. Considering these results, the 45° angle and the grey background were selected. After the improvement of the photographs, 11 food items were excluded following critical assessment by the research team. Additionally, considering suggestions from the evaluation panel, photos of 18 variants of original food items were taken (e.g., main dishes with and without salad, whole and sliced fruit). As such, a total of 62 photographs proceeded to the next evaluation phase.

3.1.2. Phase 2 of External Evaluation of the Photographs

A total of 97 photographs (62 from phase 1 and 35 photographed only in the second round of photographs production) were evaluated. In this phase, 24 food items (24.7%) were misidentified and/or misdescribed by at least one person from any of the panels. Most food items obtained a median classification of 5.0 on the perceived ease of food item recognition

374 scale; however, 20 (20.6%) were classified with a median below 4.5 by at least one of the
375 panels. Regarding portions perception, the great majority of food items (94.8%) were
376 classified as medium-sized, whereas in four photographs, portions were perceived as large,
377 indicating that the displayed images could cause an increased perception of size.
378 To assess portion size quantification, the ratio between each individual estimation and the
379 respective real portion size, both in grams, was calculated (hereafter referred to as
380 "estimation ratio"). In nearly two-thirds of food items (62.7%), the estimation ratios were
381 between 0.75 and 1.25 (i.e., under 25% error). In general, the lay panel showed higher
382 variability of estimates, and tended to under-estimate more than the experts. For 17.5% of
383 food items, the distributions of estimates were skewed towards under- ($P75_{estimation\ ratio}<1$) or
384 over-estimation ($P25_{estimation\ ratio}>1$) in both panels.
385 In light of these results, final improvements were performed in 51 photographs, of which
386 15 were later excluded following the research team's critical assessment. Also, considering
387 the commentaries from panel members, three new food items were added. As such, a final
388 pool of 85 photographs for the development of the first version of the NUTLY was obtained.
389

390 3.2. Validation Study

391 *3.2.1. External Evaluation of the First Version of the NUTLY Instrument*

392 Comparing both panels, the mean frequency of correct answers was significantly higher in
393 the experts' panel for both parameters (energy: 81.2% vs. 73.5%, p=0.030; sodium: 83.7%
394 vs. 76.3%, p=0.038). Negative correlations were observed between the perceived level of

difficulty and the frequency of right answers (ρ=-0.71, p<0.001 for energy; ρ=-0.49, p=0.020 for sodium) (**Supplementary Figure 6**).

CVI was ≥ 0.75 in 90.9% and 81.8% of energy and sodium combinations, respectively. Mean relevance scores were 3.3 for the energy parameter and 3.1 for the sodium parameter. All energy combinations presented a relevance score ≥ 3.0, whereas five sodium combinations obtained a mean relevance score below this value. Combinations with worse relevance scores were critically assessed by the research team, and two additional combinations with slight modifications were included in the instrument, for a total of 79 combinations, in order to test for the possibility of better outcomes in the following validation steps.

3.2.2. Pre-test of the NUTLY Questionnaire

Participants of the pre-test reported that, in general, the questionnaire was well structured, clear and understandable. Slight modifications in the phrasing of the initial instructions of the instrument and of specific questions were suggested by the participants and thus were applied.

3.2.3. Application of the NUTLY Questionnaire

3.2.3.1. Participants' Characteristics

In the academic community of the University of Porto, from a total of 43066 persons who received the mass email, 407 responded (391 accessing the instrument directly via the email and 16 responding in person on the premises of the faculties). A total of 18 participants were enrolled in the senior centre.

All the participants responding in person, either in the computers made available in faculties (n=16) or in the senior centre (n=18) completed the NUTLY. Among the remaining 391 who participated from their own devices, a total of 295 answered to the 79 NUTLY combinations and 96 did not complete the assessment; the latter answered to a median (percentile 25, percentile 75) of 24 (10, 37) combinations. Among the participants answering in their own devices, there were no significant differences between those who completed the NUTLY and those who answered only partially regarding gender (women: 73.6 vs. 72.9, p=0.363) or education (higher education: 68.8% vs. 60.4%, p=0.365), but those not completing the NUTLY were significantly younger (median age: 29 vs. 23 years, p<0.001) and were more likely to use a mobile phone to answer the questionnaire (21.0% vs. 36.8, p=0.008).

Overall, 329 participants who completed the NUTLY questionnaires were considered for data analysis; although 22 did not complete the METER and/or the NVS. The sociodemographic and health characteristics of the participants are shown in **Table 1**. Participants were aged between 18 and 85 years and were mostly women (72.6%), Portuguese (94.2%), with higher education in fields unrelated to health (39.8%) and self-reported normal weight (63.5%). Dyslipidaemia and high blood pressure were the most frequently reported NCDs. Most participants used a computer to fill in the questionnaire (78.1%), and a smaller part used smartphones (18.8%) or tablets (3.0%).

3.2.3.2. Item Reduction: the Final NUTLY Instrument

Two latent trait models were tested: one with a single latent variable and another with two latent variables. The former had the lowest Bayesian Information Criterion (BIC), indicating

439 that the NUTLY scale is unidimensional. A total of 29 combinations showed negative
440 standardized factor loadings (std.z1) or positive but small (0<std.z1<0.2) **(Supplementary**
441 **Table 2)**. Of those, 21 combinations were excluded after critical assessment by the research
442 team. Additionally, 10 combinations with higher associations (std.z1>0.2) were excluded
443 based on other criteria, namely: (1) low relevance considering the overall instrument's
444 purpose (n=4); (2) redundancy in relation to other combinations, in which the one with the
445 lowest performance was excluded (n=3); and (3) presence of culinary preparations in which
446 the variability in recipes was perceived as potentially higher than the difference in the
447 nutritional contents of the three options (n=3). Therefore, the final version of the NUTLY
448 instrument was obtained, and included 48 combinations: 33 for the energy parameter and
449 15 for sodium. The final instrument showed a high level of internal consistency, with a
450 Cronbach's alpha of 0.74 (95%CI: 0.70-0.78).
451 The final NUTLY instrument is displayed in **Appendix 1**, being also available in electronic
452 format through the following link: https://s.up.pt/uwlb. All resources produced in the
453 context of this instrument, namely the photographs and audio files, can also be made
454 available upon request.
455
456 *3.2.3.3. Construct Validity*
457 The NUTLY scores obtained by the participants varied between 13 and 48 on a scale from 0
458 to 48. These were highest among nutritionists (mean ± standard deviation (SD): 39.9 ± 4.4),
459 followed by those with health academic training (38.5 ± 4.1), without health academic
460 training (37.8 ± 4.4), and continued declining with decreasing education levels (p<0.001)

(**Figure 2**). The correlation between NUTLY and NVS scores was positive and statistically significant (rho=0.19; p=0.001). Considering METER's two sub-scales, correlations with the NUTLY score were both positive, but only one achieved statistical significance (words: rho=0.11; p=0.051; non-words: rho=0.14; p=0.016). Those with adequate literacy according to the NVS showed higher NUTLY scores (37.9 ± 4.3 vs. 33.9 ± 6.9, p<0.001). The same was found for METER but did not reach statistical significance (37.5 ± 4.7 vs. 36.9 ± 5.0, p=0.330).

3.2.3.4. Associations between NUTLY scores and other participants' characteristics

Table 2 shows the mean NUTLY scores according to sociodemographic, diet and health-related characteristics of the participants. Women and those who were responsible for buying or cooking foods for the household presented higher NUTLY scores. Regarding age, mean NUTLY scores peaked between the ages of 36 and 45 years, after which they gradually declined (**Supplementary Figure 7**). Participants which reported having had a previous diagnosis of high blood pressure, diabetes/pre-diabetes or cardiovascular disease had significantly lower NUTLY scores. No differences were found considering nationality, body mass index, history of overweight/obesity or dyslipidaemia, importance given to the nutritional content of the foods consumed, history of consultation with a nutritionist and device used to fulfil the questionnaire.

3.2.3.5. Test-retest reliability

From a total of 208 participants who authorised a second contact, 158 (76.0%) underwent a follow-up evaluation. The intraclass correlation between the two evaluations was 0.77

(95%CI: 0.70-0.83). The mean time spent filling out the instrument with 79 combinations was 7 minutes and 18 seconds (SD: 7 minutes and 17 seconds).

3.2.3.6. Cut-off definition, sensitivity and specificity

The Area Under the ROC Curve was 0.97. The optimum cut-off value according to the Youden index criterion was 34.5. Taking this into account, nutrition literacy levels according to the NUTLY score were dichotomized into adequate (overall score ≥ 35 correct answers out of 48 combinations) or inadequate (overall score < 35 correct answers out of 48 combinations). With this cut-off value, NUTLY showed high sensitivity (89.3%) and specificity (93.7%).

4. Discussion

We developed and validated a tool for measuring nutrition literacy in adults residing in Portugal. To our knowledge, this is the first instrument based on photographs, which proposes to directly assess the practical ability to distinguish and select foods with different nutritional characteristics. The final version of NUTLY showed to be a valid and reliable tool, with high sensitivity and specificity in the detection of individuals with limited nutrition literacy.

The comparison of latent trait models showed that the NUTLY instrument is structured in such a way that it possesses a single underlying dimension that accounts for the relationships between the items, thus enabling the construction of a unique score. According to our results, the sub-scores obtained in the energy and sodium parameters were significantly correlated, suggesting that individuals who showed higher levels of knowledge regarding the energy contents of foods also did so for sodium (and vice versa), which is in line with the unidimensionality found for this instrument. For this reason, the instrument should be considered as a whole, with a single scoring. NUTLY obtained a good level of internal consistency, in accordance with recommendations on the quality criteria for the development and evaluation of health status questionnaires (Terwee, et al., 2007), meaning that the items consistently measure the same underlying concept.

To assess construct validity, we opted to use concurrent validation, that is the comparison with other validated instruments ("ScienceDirect Topics - Concurrent Validity: an overview," 2023), complemented with testing the hypothesis that nutrition literacy should be higher in nutritionists, followed by health professionals, people with higher education in areas not

related to health, and people with progressively lower levels of formal education. We used both strategies in an attempt to overcome the multiple proposed definitions for the underlying concept (Vettori, et al., 2019) as well as the diverse and restrictive scopes of the existing instruments (Carbone & Gibbs, 2013). NUTLY scores were positively associated with education level, as well as the existence of academic training in the areas of health and nutrition, thus confirming the latter hypothesis. Many studies have demonstrated an association between education level and both health and nutrition literacy (Spronk, et al., 2014; Stormacq, Van den Broucke, & Wosinski, 2019). However, by introducing the academic training in health or nutrition, we showed that the instrument measures more than just educational attainment. Our instrument also showed a significant association with the NVS, which explores important features of the nutrition literacy concept related to the comprehension and interpretation of a nutrition label, thus reinforcing the validity of its construct. However, the same was not true regarding the association with the METER instrument. Although individuals with adequate literacy according to METER showed higher mean NUTLY scores than those with inadequate levels, differences were not statistically significant. This can be explained by the fact that the METER is an instrument that measures health literacy, and not specifically nutrition literacy. Although these concepts are interrelated, nutrition literacy is a more limited component of a broader spectrum that is health literacy. This means that a person with an adequate level of health literacy does not necessarily have a good level of nutrition literacy. Another possible explanation is that METER strongly relies on words, whereas NUTLY is based mostly on photographs. Therefore, less literate people may have low METER scores due to their inability to

recognize terms and non-terms and still have good levels of practical knowledge on the selection of nutritionally balanced foods. In our sample, from those with the lowest education level (up to 4 years of formal education), most had inadequate health literacy according to METER, but adequate nutrition literacy levels according to NUTLY. This reflects the differences in the competencies addressed by both instruments (vocabulary knowledge vs. practical ability). Even though our initial intention was to study the correlation between the NUTLY and a validated health literacy instrument, based on the results found, we concluded that using the METER to compare with NUTLY may not have been the best option, mainly due to the very distinct underlying competencies that each tool measures.

NUTLY scores were significantly higher among women, which is similar to what was found in other studies (Al Tell, Natour, Alshawish, & Badrasawi, 2023; Bonaccio, et al., 2013; Hendrie, Coveney, & Cox, 2008; Koch, et al., 2021; Parmenter, Waller, & Wardle, 2000; Worsley, Wang, Byrne, & Yeatman, 2014). Women tend to be more interested in topics related to food or diet than men, mainly due to a greater concern with body image, weight control and overall health (Bibiloni, Coll, Pich, Pons, & Tur, 2017; Forrester-Knauss & Zemp Stutz, 2012; Grogan, 2006; Yaemsiri, Slining, & Agarwal, 2011), which may lead them to seek more information on nutrition. Furthermore, it is in this group that responsibility for buying food and cooking meals usually falls. In our sample, women were more frequently responsible for cooking (68.8% vs. 38.1%, $p<0.001$) and buying foods for the household (73.0% vs. 51.2%, $p=0.001$) than men. Mean NUTLY scores were higher among those responsible for these tasks, which was expected, since they allow for the development of

food-related capabilities, such as the perception of the ingredients used in culinary preparations or the estimation of portions.

Although NUTLY scores tended to increase with higher levels of importance given to the nutritional composition of the foods consumed, differences did not reach statistical significance. This can be explained by the lower number of people answering “not important” (n=39), and by the fact that there are other determinants of food consumption that may be considered by the respondents as important competitors when selecting foods, such as price, organoleptic properties, convenience, among others (Chen & Antonelli, 2020).

Mean NUTLY scores peaked between the ages of 36 and 45 years, after which they gradually declined. This was consistent with the results found in one study using the portuguese version of the NVS, which estimated the prevalence and sociodemographic correlates of limited health literacy in a nationally-representative sample (Paiva, et al., 2017). In this study, people aged between 35 and 44 years showed the highest likelihood of adequate literacy levels, which then decreased with age. In other countries, although similar findings were observed in previous works (Hendrie, et al., 2008; Parmenter, et al., 2000), the evidence on the relationship between nutrition literacy and age has been contradictory, with studies reporting on either positive or negative associations (M. Dickson-Spillmann & Siegrist, 2011; Maria Dickson-Spillmann, Siegrist, & Keller, 2011; Koch, et al., 2021). This could be explained by the diversity of definitions, methodologies and measurement tools used, since distinct constructs of nutrition literacy have been adopted by different authors and, therefore, there is no gold standard for its assessment. Specificities in the

582 characteristics of the studied populations may also contribute to variations regarding the
583 association of nutrition literacy and socio-demographic characteristics, particularly age.
584 Lower NUTLY scores were obtained by participants reporting a previous diagnosis of high
585 blood pressure, diabetes/pre-diabetes or cardiovascular disease. Regarding body mass
586 index, although not statistically significant, normal weighted participants were the ones
587 who scored the highest, when compared with both under- and overweight. Previous studies
588 have shown that nutrition literacy skills are useful to understand key concepts and adopt
589 healthier eating habits (*e.g.,* reading nutrition information, determining portion sizes and
590 recognising healthier food choices) (Spronk, et al., 2014). As such, the lack of these skills
591 among some participants might have been associated with a higher susceptibility to
592 unhealthy eating habits and associated health outcomes such as excess weight gain and
593 development of NCDs. Further research on this association might enlighten about future
594 nutrition education interventions for health promotion using the NUTLY instrument.
595 Test-retest reliability was high, considering the recommended value of at least 0.70
596 (Nunnally & Bernstein, 1994; Terwee, et al., 2007). The period between the two evaluations
597 was relatively long (between one and two months), which allowed for the minimization of
598 learning effects (Salthouse & Tucker-Drob, 2008).
599 One of the strengths of this study was the use of multiple and diverse external evaluation
600 panels throughout the whole process of development and validation of this instrument,
601 complemented with continuous critical discussion among all team members. The positive
602 results obtained were achieved through good construct definition, as well as the
603 production, evaluation and improvement of photographs and combinations using a

stepwise, standardized process. Floor or ceiling effects (which are considered to be present if more than 15% of respondents achieved the lowest or highest possible score, respectively (McHorney & Tarlov, 1995; Terwee, et al., 2007)) did not occur. Although the NUTLY score varies between 0 and 48, an individual responding randomly would have a 33.3% probability of selecting the correct options, which would result in a score of 16. In our sample, none of the participants obtained 0, only one obtained a score lower than 16, and one achieved the maximum of 48. This indicates the existence of items across the entire spectrum of difficulty, including the lower or upper end of the scale, which benefits the instrument's content validity and the reliability of the data analysis(Liu, 2018; Terwee, et al., 2007). Regarding the time of completion, in the second evaluation (in which only the NUTLY was assessed), it was possible to observe that participants spent on average 9.1 seconds per combination. As such, we can estimate that the average time to complete the final NUTLY with 48 combinations should be around seven minutes and 18 seconds. Although this information is related to the second assessment and may be underestimated in relation to the time needed to answer when the NUTLY instrument is applied for the first time, it is plausible to consider it as a brief, easy to apply and dynamic instrument, with potential to generate high participants' adherence in future studies.

~~There are a few nutrition literacy measuring tools already validated for the Portuguese population. The General Nutrition Knowledge Questionnaire (GNKQ) (Almeida-Souza, 2009; Ferro-Lebres, et al., 2014) addresses theoretical knowledge about nutritional facts and recommendations and the Nutrition Literacy Assessment Instrument (NLAI) evaluates knowledge on nutrients, portions of food, groups of the Portuguese Food Wheel and food~~

~~labels (Azevedo, et al., 2019). A Portuguese version of the NVS was also validated (Martins & Andrade, 2014) and was used in this validation study. Although being officially described as a health literacy screening tool, it focuses on reading and interpretation of a nutrition label, therefore fitting into the nutrition literacy concept (Vettori, et al., 2019). Like other instruments, these present an essentially theoretical perspective, also being of limited application in individuals with low levels of general literacy and/or numeracy.~~ Although the photograph-based feature overcomes limitations of previous instruments, it was also one of the main challenges in the development of NUTLY, since photographs are somehow liable to have distinct interpretations. One of the objectives of the different included panels was to minimize the ambiguity of the photographs, and to maximize the ease of recognition of the foods by people of different genders, age groups and levels of education, using their feedback to improve the photographs and combinations used. Nevertheless, in practice, it is impossible to guarantee that all potential biases related to the interpretation of the photographs have been eliminated. It is also worth mentioning that it would be relevant to conduct a confirmatory study after performing item reduction on the instrument. Although the order and content of the combinations presented were not modified, the reduction in the NUTLY's length may justify the relevance of a study to confirm the psychometric properties of the final NUTLY version. Data collection for this validation study was conducted in the northern region of Portugal, which could influence the results mainly due to the regional diversity in the types of foods and dishes consumed. However, during the development of the photographs, an effort was made in order to include foods and culinary preparations ubiquitous throughout the country. In addition, people from different

geographic regions were included in the panels, in order to ensure that the foods photographed were easily recognizable by all dwellers in general. The academic community from the University of Porto includes people from diverse geographic origins as well, although we do not have that information in our specific sample. Also, no differences in NUTLY scores were found between Portuguese and non-Portuguese adults. There were no statistically significant differences between these two groups regarding gender, age, education, training in health and training in nutrition. In addition, the instrument obtained similar internal consistency and test-retest reliability values in both groups. This may reflect that the NUTLY implies a lower language barrier compared to other instruments and the selection of foods that are easily recognizable, even for individuals that have been living in Portugal for shorter periods. For this reason, we opted to include all residents in Portugal, regardless of their nationality, that way allowing for a broader inclusion of people. In addition, the fact that no differences were found in NUTLY scores of participants using computers, tablets or smartphones, suggests that the instrument may work similarly in different electronic devices. In fact, the NUTLY was developed to be compatible with different operating systems, as well as typologies, sizes and characteristics of screens, maintaining the quality of the photographs, the structure of the questionnaire and the resources presented (in particular, the audio instructions). This effort is reassuring regarding the possibility of this instrument being used in different types of devices. However, the proportion of participants not completing the NUTLY was higher among those using mobile phones, suggesting potential logistical constraints that may be removed by optimizing the instrument for use in different devices. Above all, the findings of this study

indicate that the standardized methodology used in the development of the NUTLY allows for its reproducibility, and cultural adaptation to other countries is therefore possible.

5. Conclusions

The NUTLY is a valid tool for measuring nutrition literacy in adults living in Portugal. This instrument showed good internal consistency and test-retest reliability, differentiating individuals based on education level as well as academic training in health and nutrition. NUTLY also showed high sensitivity and specificity in the detection of persons with adequate or inadequate nutrition literacy levels. A confirmatory study of the psychometric properties of the final NUTLY version would be important for strengthening the validity of this instrument.

This innovative and easy-to-use tool has the potential to be applied in the measurement of nutrition literacy across all education levels, as well as adapted to other cultures and specific population groups. In the future, this instrument could be used as part of a health surveillance system, or as a tool to assess the success of nutrition education interventions.

Funding:

This study was financed through national funding from the Foundation for Science and Technology—FCT (Portuguese Ministry of Science, Technology and Higher Education), under the projects UIDB/04750/2020, LA/P/0064/2020 and EXPL/SAU-NUT/1508/2021.

Ethical Statement:

Informed consent was obtained from all individual participants included in the study. All necessary measures were taken in order to ensure full compliance with the data protection legislation in place and the recommendations by the Data Protection Officer of the Institute of Public Health of the University of Porto. All procedures were performed in accordance with the 1964 Helsinki declaration and its later amendments. This study was approved by the Ethics Committee of the Institute of Public Health of the University of Porto, Portugal (CE19110).

Authors' contributions:

NL, PP, MS, SS, and GA designed the study. NL, GA, and PP supervised the implementation of the study. SS was responsible for the field work. MS performed data analysis. All authors critically participated in the interpretation of the results and the decision-making process throughout all phases of development and validation. SS drafted the manuscript. All authors critically revised and gave final approval of the manuscript submitted for publication.

Declarations of interest: none.

Data availability: data will be made available on request.

References

Al Tell, M., Natour, N., Alshawish, E., & Badrasawi, M. (2023). The relationship between nutrition literacy and nutrition information seeking attitudes and healthy eating patterns among a group of palestinians. *BMC Public Health, 23*, 165.

Almeida-Souza, J. (2009). *Conhecimentos Nutricionais - Reprodução e Validação do Questionário [Master's Thesis].* Faculdade de Medicina e Instituto Ciências Biomédicas Abel Salazar, Universidade do Porto.

Azevedo, J., Padrão, P., Gregório, M. J., Almeida, C., Moutinho, N., Lien, N., & Barros, R. (2019). A Web-Based Gamification Program to Improve Nutrition Literacy in Families of 3- to 5-Year-Old Children: The Nutriscience Project. *J Nutr Educ Behav, 51*, 326-334.

Bibiloni, M. d. M., Coll, J. L., Pich, J., Pons, A., & Tur, J. A. (2017). Body image satisfaction and weight concerns among a Mediterranean adult population. *BMC Public Health, 17*, 39.

Bonaccio, M., Di Castelnuovo, A., Costanzo, S., De Lucia, F., Olivieri, M., Donati, M. B., de Gaetano, G., Iacoviello, L., & Bonanni, A. (2013). Nutrition knowledge is associated with higher adherence to Mediterranean diet and lower prevalence of obesity. Results from the Moli-sani study. *Appetite, 68*, 139-146.

Carbone, E. T., & Gibbs, H. D. (2013). Measuring Nutrition Literacy: Problems and Potential Solutions. *Journal of Nutritional Disorders & Therapy, 3*, 1-2.

Chen, P. J., & Antonelli, M. (2020). Conceptual Models of Food Choice: Influential Factors Related to Foods, Individual Differences, and Society. *Foods, 9*.

Dickson-Spillmann, M., & Siegrist, M. (2011). Consumers' knowledge of healthy diets and its correlation with dietary behaviour. *J Hum Nutr Diet, 24*, 54-60.

Dickson-Spillmann, M., Siegrist, M., & Keller, C. (2011). Development and validation of a short, consumer-oriented nutrition knowledge questionnaire. *Appetite, 56*, 617-620.

Elmskini, F. Z., Bouh, A., Labyad, A., Elghoulam, N., Iraqi, H., Mehdad, S., Madkour, A., Moufid, A., Aabi, M., Boutayeb, S., Taghzouti, K., Aguenaou, H., & Benaich, S. (2024). Increased nutrition knowledge and adherence to the Mediterranean diet are associated with lower body mass index and better self-rated general health among university students. *Human Nutrition & Metabolism, 35*, 200240.

EuroFIR. (2008). Report on collection of rules on use of recipe calculation procedures including the use of yield and retention factors for imputing nutrient values for composite foods In.

Ferro-Lebres, V., Moreira, P., & Ribeiro, J. C. (2014). Adaptation, Update and Validation of the General Nutrition Questionnaire in a Portuguese Adolescent Sample. *Ecol Food Nutr, 53*, 528-542.

Forrester-Knauss, C., & Zemp Stutz, E. (2012). Gender differences in disordered eating and weight dissatisfaction in Swiss adults: Which factors matter? *BMC Public Health, 12*, 809.

Goios, A., Martins, M. L., Oliveira, A. C., Afonso, C., & Amaral, T. (2016). *Pesos e Porções de Alimentos. 2nd ed.* Porto, Portugal: U Porto Press.

753 Grogan, S. (2006). Body image and health: contemporary perspectives. *J Health Psychol, 11*,
754 523-530.
755 Grunert, K. G., Wills, J. M., & Fernández-Celemín, L. (2010). Nutrition knowledge, and use
756 and understanding of nutrition information on food labels among consumers in the
757 UK. *Appetite, 55*, 177-189.
758 Hendrie, G. A., Coveney, J., & Cox, D. (2008). Exploring nutrition knowledge and the
759 demographic variation in knowledge levels in an Australian community sample.
760 *Public Health Nutrition, 11*, 1365-1371.
761 Instituto Nacional de Saúde Doutor Ricardo Jorge. (2021). Tabela da Composição de
762 Alimentos, v 5.0. In. Lisbon, Portugal: INSA.
763 Jansen, T., Rademakers, J., Waverijn, G., Verheij, R., Osborne, R., & Heijmans, M. (2018).
764 The role of health literacy in explaining the association between educational
765 attainment and the use of out-of-hours primary care services in chronically ill
766 people: a survey study. *BMC Health Services Research, 18*, 394.
767 Jezewska-Zychowicz, M., & Plichta, M. (2022). Diet Quality, Dieting, Attitudes and Nutrition
768 Knowledge: Their Relationship in Polish Young Adults-A Cross-Sectional Study. *Int J*
769 *Environ Res Public Health, 19*.
770 Koch, F., Hoffmann, I., & Claupein, E. (2021). Types of Nutrition Knowledge, Their Socio-
771 Demographic Determinants and Their Association With Food Consumption: Results
772 of the NEMONIT Study. *Front Nutr, 8*, 630014.
773 Liu, Q. (2018). Data Analyses with Ceiling/Floor data. In: R Project.
774 Lopes C, Torres D, Oliveira A, Severo M, Alarcão V, Guiomar S, Mota J, Teixeira P, Rodrigues
775 S, Lobato L, Magalhães V, Correia D, Carvalho C, Pizarro A, Marques A, Vilela S,
776 Oliveira L, Nicola P, Soares S, & E., R. (2017). Inquérito Alimentar Nacional e de
777 Atividade Física, IAN-AF 2015-2016: Relatório de resultados. In: Universidade do
778 Porto.
779 Martins, A. C., & Andrade, I. M. (2014). Cross-cultural adaptation and validation of the
780 portuguese version of the Newest Vital Sign. *Revista de Enfermagem Referência, 4*,
781 75-83.
782 McHorney, C. A., & Tarlov, A. R. (1995). Individual-patient monitoring in clinical practice:
783 are available health status surveys adequate? *Qual Life Res, 4*, 293-307.
784 Mohsen, H., Sacre, Y., Hanna-Wakim, L., & Hoteit, M. (2022). Nutrition and Food Literacy in
785 the MENA Region: A Review to Inform Nutrition Research and Policy Makers. *Int J*
786 *Environ Res Public Health, 19*, 10190.
787 Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory, 3rd ed.* New York: McGraw-
788 Hill.
789 Paiva, D., Silva, S., Severo, M., Ferreira, P., Santos, O., Lunet, N., & Azevedo, A. (2014). Cross-
790 cultural adaptation and validation of the health literacy assessment tool METER in
791 the Portuguese adult population. *Patient Educ Couns, 97*, 269-275.
792 Paiva, D., Silva, S., Severo, M., Moura-Ferreira, P., Lunet, N., & Azevedo, A. (2017). Limited
793 Health Literacy in Portugal Assessed with the Newest Vital Sign. *Acta Med Port, 30*,
794 861-869.
795 Parmenter, K., Waller, J., & Wardle, J. (2000). Demographic variation in nutrition knowledge
796 in England. *Health Educ Res, 15*, 163-174.

797 Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content
798 validity? Appraisal and recommendations. *Res Nurs Health, 30*, 459-467.
799 PortFIR – Grupo de Trabalho Porções. (2014). Guia Orientativo Para o Estabelecimento de
800 Porções para a Rotulagem Nutricional. In. Lisbon, Portugal: Instituto Nacional de
801 Saúde Doutor Ricardo Jorge.
802 Ratzan, S. C., & Parker, R. M. (2000). *National Library of Medicine Current Bibliographies in*
803 *Medicine: Health Literacy*. Bethesda, MD: National Institutes of Health, U.S.
804 Department of Health and Human Services.
805 Salthouse, T. A., & Tucker-Drob, E. M. (2008). Implications of short-term retest effects for
806 the interpretation of longitudinal change. *Neuropsychology, 22*, 800-811.
807 . ScienceDirect Topics - Concurrent Validity: an overview. In. (2023): Elsevier B.V.
808 Spronk, I., Kullen, C., Burdon, C., & O'Connor, H. (2014). Relationship between nutrition
809 knowledge and dietary intake. *Br J Nutr, 111*, 1713-1726.
810 Stormacq, C., Van den Broucke, S., & Wosinski, J. (2019). Does health literacy mediate the
811 relationship between socioeconomic status and health disparities? Integrative
812 review. *Health Promot Int, 34*, e1-e17.
813 Taylor, M. K., Sullivan, D. K., Ellerbeck, E. F., Gajewski, B. J., & Gibbs, H. D. (2019). Nutrition
814 literacy predicts adherence to healthy/unhealthy diet patterns in adults with a
815 nutrition-related chronic condition. *Public Health Nutr, 22*, 2157-2169.
816 Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter,
817 L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement
818 properties of health status questionnaires. *J Clin Epidemiol, 60*, 34-42.
819 van der Heide, I., Uiters, E., Sørensen, K., Röthlin, F., Pelikan, J., Rademakers, J., Boshuizen,
820 H., & consortium, o. b. o. t. E. (2016). Health literacy in Europe: the development
821 and validation of health literacy prediction models. *European Journal of Public*
822 *Health, 26*, 906-911.
823 Vettori, V., Lorini, C., Milani, C., & Bonaccorsi, G. (2019). Towards the Implementation of a
824 Conceptual Framework of Food and Nutrition Literacy: Providing Healthy Eating for
825 the Population. *Int J Environ Res Public Health, 16*.
826 Walter, S. D., & Irwig, L. M. (1988). Estimation of test error rates, disease prevalence and
827 relative risk from misclassified data: a review. *J Clin Epidemiol, 41*, 923-937.
828 World Health Organization. (2013). Health literacy - The solid facts. In. Copenhagen,
829 Denmark: World Health Organization Europe.
830 World Health Organization. (2018). Noncommunicable diseases country profiles. In.
831 Geneva, Switzerland: World Health Organization.
832 World Health Organization. (2022). Health literacy development for the prevention and
833 control of noncommunicable diseases. In. Geneva, Switzerland: World Health
834 Organization
835 Worsley, A., Wang, W. C., Byrne, S., & Yeatman, H. (2014). Different patterns of Australian
836 adults' knowledge of foods and nutrients related to metabolic disease risk. *J Nutr*
837 *Sci, 3*, e14.
838 Yaemsiri, S., Slining, M. M., & Agarwal, S. K. (2011). Perceived weight status, overweight
839 diagnosis, and weight control among US adults: the NHANES 2003–2008 Study.
840 *International Journal of Obesity, 35*, 1063-1070.

841 Zajacova, A., & Lawrence, E. M. (2018). The Relationship Between Education and Health:
842 Reducing Disparities Through a Contextual Approach. *Annu Rev Public Health, 39*,
843 273-289.

844

Table 1. Sociodemographic and health characteristics of the participants (n=329).

	n	%
Gender		
Male	85	25.8
Female	239	72.6
Prefers not to answer	5	1.5
Age (years)		
Mean ± SD = 36.7 ± 16.1		
18-25	116	35.3
26-35	65	19.8
36-45	53	16.1
46-55	44	13.4
56-65	37	11.2
>65	14	4.3
Educational level and academic training		
Basic education (up to 9 years)	38	11.6
Secondary education (12 years)	86	26.1
Higher education – unrelated to health area	131	39.8
Higher education – health area	53	16.1
Higher education – nutrition area	21	6.4
Nationality		
Portuguese	310	94.2
Non-Portuguese	19	5.8
Body Mass Index (kg/m^2)*		
Mean ± SD = 24.2 ± 4.7		
Underweight	9	2.8
Normal weight	202	63.5
Overweight	72	22.6
Obesity	35	11.0
History of NCDs		
Dyslipidaemia	74	22.5
High blood pressure	43	13.1
Diabetes or pre-diabetes	22	6.7
Cardiovascular diseases	8	2.4
Device used to answer the questionnaire		
Computer	257	78.1
Smartphone	62	18.8
Tablet	10	3.0

NCDs, non-communicable diseases; SD, standard deviation.

***Calculated based on self-reported weight and height, and classified according to the World Health Organization's cut-offs (n=318).**

Table 2. Mean NUTLY scores according to sociodemographic, diet and health-related characteristics (n=329).

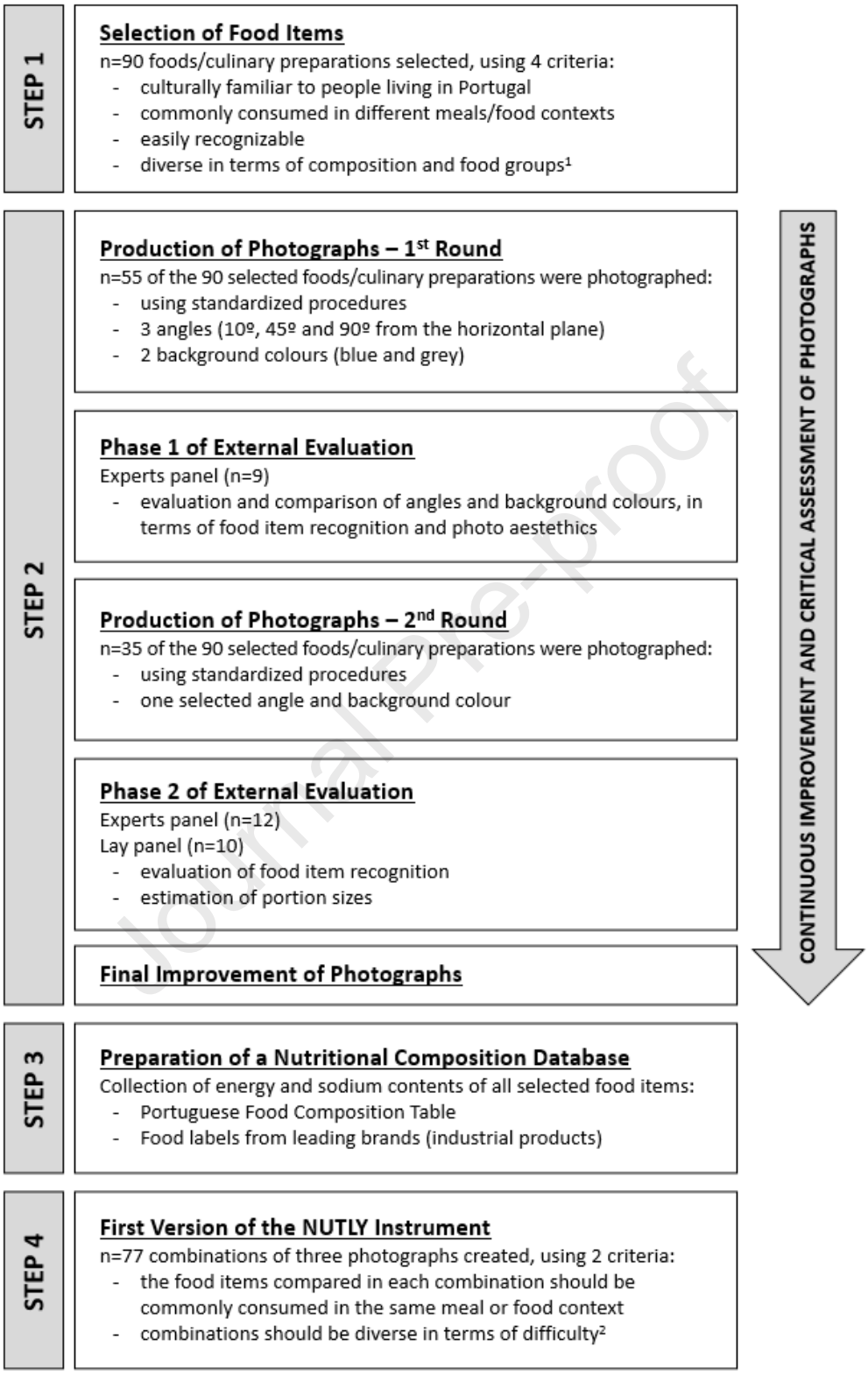
	NUTLY score		p-value
	Mean	SD	
Gender			
Male	35.7	5.7	
Female	37.6	5.0	**0.016**
Prefers not to answer	35.6	2.6	
Age (years)			
18-25	36.9	4.2	
26-35	37.6	4.4	
36-45	38.4	4.7	**0.015**
46-55	37.8	5.9	
56-65	35.5	7.1	
>65	33.6	8.2	
Nationality			
Portuguese	37.1	5.3	0.690
Non-Portuguese	36.7	4.1	
Body Mass Index (kg/m^2)*			
Underweight	34.2	5.7	
Normal weight	37.4	5.0	0.166
Overweight	36.8	5.1	
Obesity	37.0	5.2	
History of NCDs			
Dyslipidaemia			
Yes	36.2	6.2	0.097
No	37.3	4.9	
High blood pressure			
Yes	35.3	6.6	**0.022**
No	37.3	5.0	
Diabetes or pre-diabetes			
Yes	34.0	8.1	**0.004**
No	37.3	4.9	
Cardiovascular diseases			
Yes	32.5	10.4	**0.013**
No	37.2	5.0	
Responsibility of buying foods for the household			
Yes	37.7	4.9	**0.002**
No	35.7	5.5	
Responsibility of cooking foods for the household			
Yes	37.7	5.0	**0.004**
No	36.0	5.5	
Importance given to the nutritional content of foods			
Not important	36.0	4.4	
Important	36.8	5.5	0.060
Very important	38.0	4.8	
History of consultation with a nutritionist			
Yes	37.2	4.3	0.550
No	36.9	5.9	
Device used to answer the questionnaire			
Computer	37.0	5.5	
Smartphone	37.3	4.0	0.902
Tablet	37.4	5.4	

NCDs, non-communicable diseases; SD, standard deviation.
*Calculated based on self-reported weight and height, and classified according to the World Health Organization's cut-offs.

Note: Values in bold represent statistically significant differences according to t-test (2 groups) or one-way ANOVA (>2 groups) with a significance level of 0.05.

Figure 1. Flowcharts summarizing (a) the development and (b) the validation of the NUTLY instrument.

(a)



[1]Food groups: main dishes; soups; sandwiches; savoury pastries and snacks; sweet cakes and pastries; desserts; confectionery; cookies and biscuits; bread; charcuterie and cheese; and fruit.

[2]The difficulty of the comparison was defined for this purpose as the percentage variation between the highest content option (i.e., right answer) and the other two options.

(b)

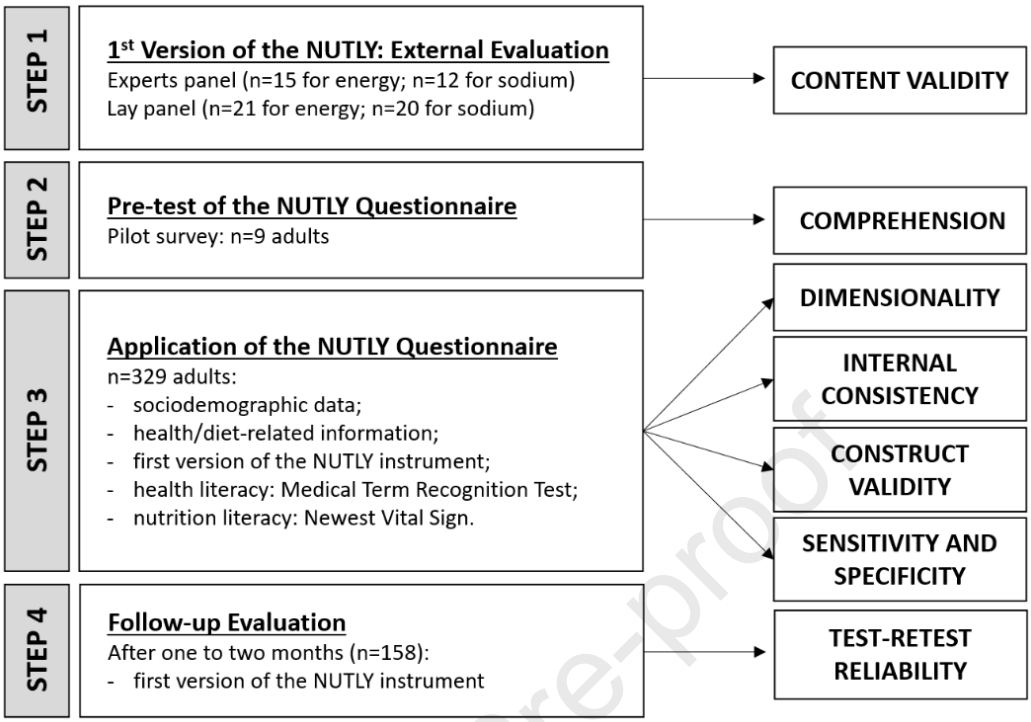
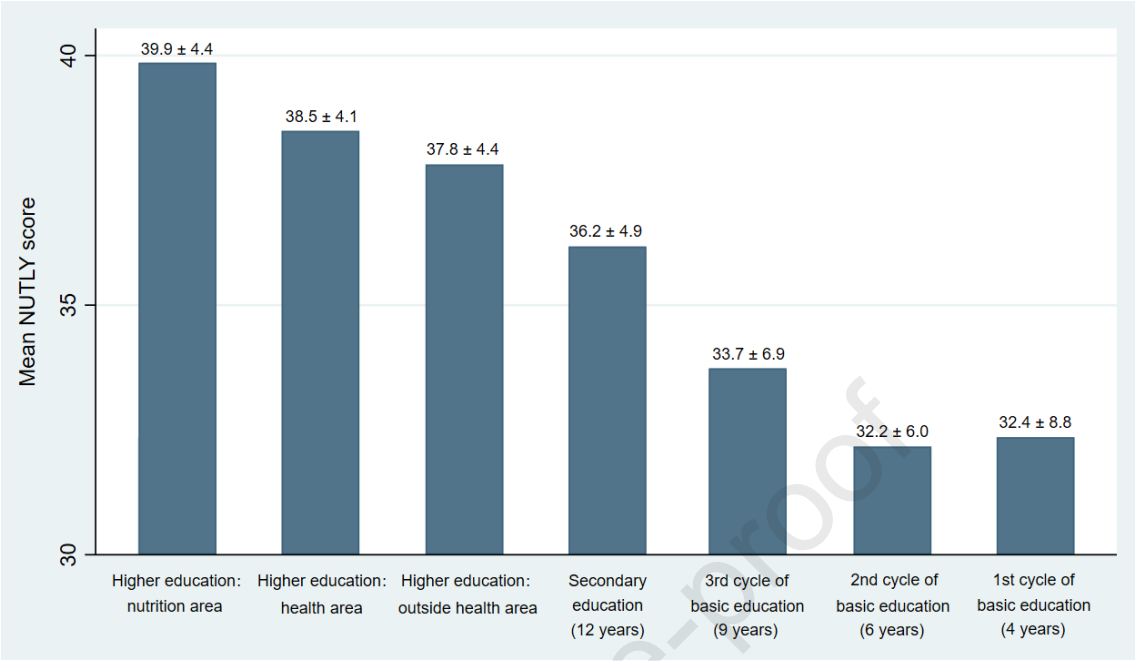


Figure 2. NUTLY scores (mean ± standard deviation) according to education level and training in health/nutrition.



Ethical Statement:

Informed consent was obtained from all individual participants included in the study. All necessary measures were taken in order to ensure full compliance with the data protection legislation in place and the recommendations by the Data Protection Officer of the Institute of Public Health of the University of Porto. All procedures were performed in accordance with the 1964 Helsinki declaration and its later amendments. This study was approved by the Ethics Committee of the Institute of Public Health of the University of Porto, Portugal (CE19110).

Funding:

This study was financed through national funding from the Foundation for Science and Technology—FCT (Portuguese Ministry of Science, Technology and Higher Education), under the projects UIDB/04750/2020, LA/P/0064/2020 and EXPL/SAU-NUT/1508/2021.

Authors' contributions:

NL, PP, MS, SS, and GA designed the study. NL, GA, and PP supervised the implementation of the study. SS was responsible for the field work. MS performed data analysis. All authors critically participated in the interpretation of the results and the decision-making process throughout all phases of development and validation. SS drafted the manuscript. All authors critically revised and gave final approval of the manuscript submitted for publication.

Declarations of interest: none.