Statistica & Applicazioni - Vol. XVII, n. 1, 2019, pp. 63-84

A NEW TOOL TO ASSESS THE CONCEPT OF FOOD QUALITY IN CHILDREN: A VALIDATION STUDY

Silvia Golia[1*]
Anna Simonetto[2*]
Buirma Malo[2]
Gianni Gilioli[2]

SUMMARY

Childhood is the most relevant period for the formation of food preferences. Political institutions, educational agencies and health professionals are quite interested in the evaluation of the actual knowledge of food quality in children because it offers the opportunity to plan educational actions aimed at promoting the health and quality of life of today's and tomorrow's citizens. In this paper, we proposed a questionnaire to measure the level of knowledge of the food quality in children. We selected, as the target population, students of lower secondary schools and we adopted a multidimensional approach to the concept of food quality that involved six dimensions. We investigated the internal construct validity and reliability of the scales developed for these six dimensions applying the Rasch model and a differential item functioning analysis. Moreover, we validated the overall scale of the knowledge of food quality through a Second-order Structural Equation Modelling within a Confirmatory Factor Analysis. The results of our study provide information useful to define and implement educational actions in food quality, in general, or specifically in one or more of food quality dimensions.

Keywords: *Rasch Model, Differential Item Functioning Analysis, Second-order Structural Equation Modelling, Confirmatory Factor Analysis.*
DOI: 10.26350/999999_000022
ISSN: 18246672 (print)

1. INTRODUCTION

In the last years, an improvement in life conditions contributed to a reduction in hunger, infant mortality and poverty, albeit the problem of undernutrition still represents a critical and relevant issue. At the same time, the share of people adopting unhealthy food styles is increasing with harmful consequences on health, e.g. increased burden of obesity and other non-communicable diseases. Undernutrition and unhealthy diets characterize the ''double burden'' of malnutrition around the world (Popkin, Corvalan

[1] Dipartimento di Economia e Management - Università di Brescia - C.da S. Chiara, 50 - 25122 BRESCIA (e-mail: silvia.golia@unibs.it).
[2] Agrofood Lab, Department of Molecular and Translational Medicine - Università di Brescia - Viale Europa, 11 - 25123 BRESCIA (e-mail: ✉ anna.simonetto@unibs.it; b.malo@studenti.unibs.it; gianni.gilioli@unibs.it).
[*] These authors contributed equally to this work.

and Grummer-Strawn, 2020). Another key issue to consider is the impact of food systems on the environment. About 40% of globe's land surface is used for agricultural purposes and food production is responsible for up to 26% of global greenhouse gas emissions (Poore and Nemecek, 2018) and 70% of freshwater use (FAO, 2017). Therefore, the need for a sustainable food production system emerges clearly, as well as for new eating habits, considered as the main drivers of food production. In such a complex system, all the actors (i.e. consumers, producers, traders, institutions) must be involved for the definition of a shared idea of food quality encompassing all the different aspects (e.g. individual, social, economic, environmental, politic) that can make it up (Willett, Rockström, Loken, Springmann, Lang, Vermeulen and Jonell, 2019).

The food quality is an emerging property widely discussed in the literature without yet agreeing to a scientifically mature definition. Food quality refers to several aspects of food (from production to consumption and waste disposal), which can vary according to the objective of the analysis (Cardello, 1995). However, there is a set of dimensions that recur in most recent studies, mainly proposed by The Food and Agriculture Organization (FAO) and World Health Organization (WHO) (Gustafson, Gutman, Leet, Drewnowski, Fanzo and Ingram, 2016; FAO and WHO, 2019, HLPE, 2017; Vågsholm, Arzoomand and Boqvist, 2020; WHO, 2015): Food Safety, Food Security, Nutritional Value, Organoleptic Quality, Environmental Sustainability and Social Responsibility. In this paper we based the define the food quality as emerging property of these six dimensions. Food Safety refers to handling, preparing and storing food in the most appropriate way to decrease the risk of foodborne illnesses. Ensuring food safety to protect public health and promote economic development remains a significant challenge in both developing and developed countries. Food security is related to the availability and the possibility of having access to a sufficient quantity of food (of appropriate quality) to satisfy nutritional requirements of people. Food security includes also the concepts of utilization, which consists in the knowledge necessary to adequately prepare and consume food, and the stability, which consists in having available food and appropriate food access over time. Every food contains specific nutrients (vitamins, proteins, fat, minerals, water etc) which constitute its nutritional value. Organoleptic quality refers to the perceptions, such as taste, aroma, colour, consistency, experienced by means of the five senses during food consumption. It plays a major role in food acceptance, influencing food preferences and choices. Environmental sustainability of food refers to the environmental impact of food production, conservation and transformation processes. Food waste also plays a key role in this context. The corporate social responsibility refers to the actions promoted by companies to promote economic, environmental and social progress and to reduce negative impacts of their activities.

Children often suffer the biggest consequences regarding non-appropriate food and nutrition choices (WHO, 2018). They are consumers to all intents and purposes because they have preferences and desires and try to satisfy them, they have a high power to influence family purchasing decisions and they manage money ever more precociously (Perez, 2009). There is a ''co-construction'' of food choice between parents and children, strongly influenced by external pressures, such as advertising

(Holsten, Deatrick, Kumanyika, Pinto-Martin and Compher, 2012). Actions at different institutional and educational levels should be taken to improve nutrition and food knowledge in children, as strongly recommended by relevant authorities dealing with food systems and nutrition (e.g. FAO, WHO). Educative actions on food and nutrition could prevent obesity in children and adolescents (Flodmark, Marcus and Britton, 2006). A key element for the success of these educative interventions is the ability of evaluating the actual knowledge of students in terms of food quality, in order to identify knowledge gaps and whether the proposed actions can effectively fill them. In this paper, we focused the attention on children of the lower secondary school. Being one of the most sensitive periods for the definition of food preferences, nutrition education programs should be target to this age.

To assess whether and how children know the concept of food quality according to the six dimensions previously introduced, a specific questionnaire was developed (see Appendix A). This study shows the outcomes of the validation procedure of the scales referring to the knowledge of the six specific dimensions of food quality and to the overall knowledge of food quality. Therefore, the first step of the analysis consisted in investigating the internal construct validity and reliability of the scales for the six dimensions. Following the *Standards for educational and psychological testing* (American Educational Research Association, American Psychological Association and National Council on Measurement in Education, 2014), 'validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests', so 'a process of validation involves accumulating relevant evidence to provide a sound scientific basis for the proposed score interpretation'. Various sources of evidence can be used covering different aspects of the validity, e.g. test content, response processes, internal structure, relationship to other variables, and consequences of testing. Regarding internal structure, the evidences include unidimensionality and absence of differential item functioning. One of the models that can be applied to address several aspects related to scale validation is the Rasch Model (Bond and Fox, 2015). The Rasch Model (RM) is a measurement model that, under the hypotheses of unidimensionality and local independency, converts raw scores into objective measures. Therefore, in this study we used the RM to investigate the validity of the proposed scales and drew information regarding the knowledge of the six different aspects that characterize the food quality. In the second step of the analysis the overall knowledge of food quality was validated through a multidimensional approach. To test the reliability and validity of the overall scale we performed an analysis based on a Second-order Structural Equation Modelling (SEM) within a Confirmatory Factor Analysis (CFA), according to Muthén approach (Muthén, 1984; Kaplan, 2000).

The results of this study could provide information useful to define and implement educational actions on food quality, both on the general construct, and specifically on one or more dimensions.

The paper is organized as follows. Section 2 reports the description of the survey and the sample of involved students. Section 3 contains a brief description of the applied statistical tools, whose results are shown in Section 4. Conclusions follow in Section 5.

2. DESCRIPTION OF THE STUDY

2.1 *The questionnaire*

An assessment tool was developed to evaluate the knowledge of food quality owned by the children. Based on the suggestion of pedagogical experts, we chose a written questionnaire for self-administration. The assessment tool was composed by three parts: 1) socio-demographic data; 2) an open question in which students were asked to describe what is meant by food quality; 3) a questionnaire to evaluate student's knowledge in terms of multidimensional food quality. In the first section, students were asked if they had ever had any previous food education experience (both school and extra-school). The possible answer was 'yes' or 'no'. The questionnaire was composed by 60 true/false items (ten items for each dimension of food quality).

To the best of our knowledge, there are no validated scales to analyse the children knowledge on the six dimensions of food quality. For this reason, we listed some key aspects for each dimension and we formulated a set of questions to evaluate whether the children knew them or not. Key aspects have been selected from the definitions shared in the literature and, where present, from the regulations. For food safety (*FSaf*) we mainly referred to FAO and WHO recommendations for children (WHO, 2006; WHO and FAO, 2012). Items investigating food security (*FSec*) were formulated based on the definition proposed by FAO (1996). In 2018, WHO published an interesting paper on effective actions for improving adolescent nutrition and we used it to define the set of items related to nutritional value (*Nutr*). For organoleptic quality (*OrgQ*) we referred to several studies investigating the role of sensory properties and the food pleasantness on food preference, some of them specifically aimed at children (Forde, 2018; Piqueras-Fiszman, Alcaide, Roura and Spence, 2012; Marzi, 2008). The section on environmental sustainability (*Sust*) is based on FAO reports (FAO, 2011; 2017; 2018). Social responsibility (*SocR*) includes aspects drawn from documents published by the European Commission (2001) and the International Organization for Standardization (ISO, 2004).

For each statement, there were three possible answers: *True, False,* and *Don't know*. Considering that the objective of the analysis was the evaluation of the actual knowledge of the children about food quality, the students were encouraged to answer only in the case of effective knowledge of the topic, not to try to guess. For this reason, in model development we considered the *Don't know* answer as wrong, because it is a sign of lack of knowledge of the topic.

2.2 *The survey*

In this study, we focused our attention on students of the lower secondary school (11-14 years old). The selection of the schools was based on a territorial criterion. We divided the Brescia's province (Italy) into 6 areas (capital city and 5 homogeneous areas by population size) and we asked the provincial school authority to indicate, for each area, a school with at least six classes. In five schools, six classes per institute were involved, whereas in a school, due to administrative issues, only four classes were interviewed. In total, 696 students were involved. Data were collected from November 2018 to January 2019. The age of respondents varies from 10 to 16 years old (12.49 ± 1.09 years). Table 1 reports the distribution of respondents by school and class attended.

TABLE 1. - *Distribution of respondents by school and attended class*

School	*Number of respondents*	*Class attended*	*Number of respondents*
School 1	128 (18.39%)	*I*	240 (34.48%)
School 2	138 (19.83%)	*II*	220 (31.61%)
School 3	123 (17.67%)	*III*	236 (33.91%)
School 4	125 (17.96%)	*Total*	696
School 5	128 (18.39%)		
School 6	54 (7.76%)		
Total	696		

Table 2 reports the percentage of correct and wrong responses to the items as well as the percentage of Don't know and missing answers.

TABLE 2. - *Percentage of correct (C), wrong (W), Don't know (DN) and missing answers (NA) according to the investigated food quality dimensions and attended class (I, II, III)*

Food quality dimension	*I*				*II*				*III*			
	C	*W*	*DN*	*NA*	*C*	*W*	*DN*	*NA*	*C*	*W*	*DN*	*NA*
FSaf	45.8	21.7	30.4	2.1	45.5	19.1	35.0	0.5	47.5	19.5	32.2	0.8
FSec	71.2	15.4	11.7	1.7	68.2	15.9	15.9	0.0	72.9	11.9	15.3	0.0
Nutr	33.3	40.8	23.8	2.1	46.4	42.3	10.9	0.5	58.5	28.8	12.3	0.4
OrgQ	40.0	28.7	29.2	2.1	47.3	14.1	38.2	0.5	53.4	15.7	29.7	1.3
Sust	39.2	14.6	44.6	1.7	40.5	15.0	43.6	0.9	42.4	9.3	46.2	2.1
SocR	77.9	6.2	13.3	2.5	83.2	2.7	13.6	0.5	87.7	2.5	8.1	1.7

From Table 2 it is possible to observe that there is a high heterogeneity in the knowledge of the six dimensions of food quality. *FSec* and *SocR* are the ones with higher percentages of correct answers for each attended class, whereas *Sust* has the highest percentages of *Don't know*, almost 50% for each attended class.

3. STATISTICAL METHODS

In this section, the models used for the validation study are presented. In order to validate the six sub-scales, which describe the six dimensions of the food quality previously introduced, we used the RM and the DIF analysis. To validate the overall knowledge on food dimension, we performed a CFA based on a second-order SEM.

3.1 *The Rasch Model*

The Rasch Model (RM) (Rasch, 1960) is a measurement model which converts raw scores into linear and reproducible measurements. Its peculiar characteristics are separable person and item parameters, sufficient statistics for the parameters and conjoint additivity. The RM's hypotheses are unidimensionality and local independence. Unidimensionality means that all the items must measure the same latent trait, that is the one of interest. Local independence means that, conditional to the latent trait, the response to an item in the test is independent to the responses given to the remaining items. If the data fit the model, the corresponding measures are objective and their measurement unit is the logit, which is an interval scale (Bond and Fox, 2015; Wright and Master, 1982).

The RM assumes that the probability of response to an item given by a person depends only on the difficulty of the item and the ability of the person according to the following equation:

$$P(X_{si} = x) = \frac{exp\{x(\beta_s - \gamma_i)\}}{1 + exp\{\beta_s - \gamma_i\}}, \qquad x = 0{,}1$$

where β_s measures the ability of person s and γ_i the difficulty of item i. Both β_s and γ_i are expressed in logit.

Tools useful to evaluate the accordance between the model and the data are the Infit and Outfit Mean Square statistics (Intfit/Outfit MNSQ), the principal component analysis (PCA) without rotation on standardized residuals and the standardized residual correlation between couples of items (Bond and Fox, 2015).

Infit and Outfit MNSQ evaluate how accurately the data fit the model and they express a sensitivity to different deviations, that is Outfit statistic is more sensitive to extreme values, whereas Infit statistic is more sensitive to pattern of responses to items targeted on the subject and vice versa. Their expected value is equal to one, indicating a negligible distortion of the measurement system, whereas values less (bigger) than one indicate predictability (unpredictability and un-modelled noise).

The PCA on standardized Rasch residuals allows verifying the assumption of unidimensionality required by the RM. In Rasch analysis, the hypothesis is that there is only one dimension, caught by the model, so we expect that the residuals do not contain other significant dimensions. Therefore, we verify the absence of other dimensions by means of the eigenvalues of PCA on Rasch residuals, focusing on the first eigenvalue λ_1.

The standardized residual correlation allows one to verify the local independence hypothesis. It is computed between residuals across all subjects who responded to both items, and high positive or negative correlations indicate potentially locally dependent pairs of items.

In order to study the reliability of the test, the Greatest Lower Bound (GLB) was used. The choice to use this reliability index instead of the widely used Cronbach's Alpha, relies on the fact that Cronbach's Alpha in many situations underestimates the reliability of the test (McNeish, 2018)

The items' difficulties γ_i and persons' abilities β_s were estimated using the joint maximum likelihood estimation method (Wright and Master, 1982) under the constraint that the sum of item difficulty parameters was equal to 0.0 logits. This constraint allows giving an interpretation to the distance of the mean difficulty from zero, which is, under this constraint, the mean difficulty of the items. In fact, a mean logit score of zero indicates a well-targeted scale, without floor or ceiling effects, whereas an average ability significantly higher (lower) than zero, indicates that the test is too easy (difficult), so it should be revised. Nevertheless, if the test covers all the main facets of the latent trait under study, this discrepancy can be interpreted as the evidence that, overall, the sample of respondents owns a high (low) level of ability; in the present context it would mean that the students have a high (low) level of knowledge of the specific food construct.

3.2 *Differential Item Functioning analysis*

Differential Item Functioning (DIF) is an item characteristic. It refers to the fact that, for a given item, the expected responses to that item given by persons with the same ability, but from different identifiable groups, are not the same. In the case of dichotomously scored items, the expected response reduces to the probability of a correct response. In the usual setting of a DIF study, the considered identifiable groups are two, often referred as focal and reference groups, but it often happens in real contexts that there are more focal groups. The common approach is to perform pairwise comparisons between the reference group and each focal group, nevertheless the power to detect DIF items with this procedure is lower than the use of a single test, if available, which simultaneously compares all groups (Penfield, 2001). In literature, there are many approaches to test the presence of DIF; Penfield and Camilli (2007), among others, overview them for both dichotomous and polytomous items. The main distinction is between the Item Response Theory (IRT) and Non-IRT approach. Methods belonging to the IRT approach require, as a first step, the estimation of an IRT model, unlike the methods belonging the Non-IRT approach. In this study, we used a DIF method for dichotomous items belonging to the Non-IRT approach, that is the Mantel-Haenszel method (Mantel and Haenszel, 1959; Holland and Thayer, 1988) and its generalization (the generalized Mantel-Haenszel method) to the case of more than two groups (Somes, 1986; Penfield, 2001). The Mantel-Haenszel methods test whether an association between group membership and item response, conditionally to the total test score, exists. The null hypothesis is the absence of this association, which translates in the hypothesis of no DIF.

In general, a DIF study involves many items and this fact implies that two aspects should be considered. The first regards the possibility that the presence of one or more DIF items influences the result of testing for DIF of the other items. To overcome this potential problem, a procedure called item purification can be implemented (Magis, Beland, Tuerlinckx and De Boeck, 2010). It is an iterative procedure which consists, firstly, in testing for DIF all items one by one, assuming they are not DIF items, and then, on the basis of the results of the previous analysis, in defining a set of DIF items. If this set is empty, the procedure stops, otherwise the iterative procedure starts and the test for DIF is repeated for all items one by one, omitting the DIF items, except the one to be tested, found at the previous step. If a different set of DIF items is identified, than test for DIF is repeated, otherwise the procedure stops. The second aspect regards the presence of multiple tests, and this can be a problem, give that multiple testing could increase the probability to commit a Type I error at least once, implying the possibility to incorrectly identify non-DIF items as DIF items. In order overcome this second potential problem, it is possible to adjust for multiple comparisons. Instead of using the Bonferroni correction, we use the Benjamini-Hochberg false discovery rate, which performs better in the DIF context, according to Kim and Oshima (2013).

3.3 *SEM*

Structural Equation Model (SEM) is a modelling framework aimed at establishing and quantifying relationships between variables (Bollen, 1989). A key feature of SEM is the ability to measure ‘latent variables’ (or factors), not directly observable, by means of a set of observable and measurable variables (indicators or items). This feature has made SEM highly appreciated for the study of complex phenomena in different areas. SEM can be explored to define the structure of relationships between latent and observed variables (exploratory factor analysis) or to assess if a conceptual model, i.e. a priori hypothesized relationships between a set of variables, is consistent with empirical data (CFA).

SEM models are characterized by structural and measurement models. In the structural model the relationships between latent constructs are specified by

$$\theta = \mathbf{B}\theta + \zeta$$

In the measurement model the relationships between latent and observed variables are specified by

$$y = \Lambda^y \theta + \epsilon$$
$$x = \Lambda^x \theta + \delta$$

where θ is the $(m \times 1)$ vector of the m latent variables, **B** is the coefficient matrix for the effects of latent variables on each other, and ζ is the $(m \times 1)$ vector of disturbances, representing the errors in structural equations. y and x are the $(p \times 1)$ and $(q \times 1)$ vectors containing the measures of the p endogenous constructs, and the q exogenous indicators, respectively. Λ^y and Λ^x are the coefficient matrices for the relations of y and x with θ (called loadings). ϵ and δ are the vector of disturbances, representing the measurement errors. According to Muthén (1983), dichotomous indicators are considered as realizations of underlying continuous indicators (Golia and Simonetto, 2015).

To validate the overall knowledge of food quality (*OveK*) we defined a second-order reflective factor analysis. The first-order factors are the latent variables representing the six dimensions of food quality. Each first-order factor is associated with the ten dichotomous indicators represented by the 10 items of the questionnaire that investigate that dimension. *OveK* is the second-order whose indicators are represented by the first-order factors are indicators (Byrne, 2013).

SEM model will be tested to assess the internal consistency, i.e. measure based on the correlations between different indicators of the same latent variable, and unidimensionality of each first-order latent variables (Simonetto, 2012).

The diagnostic indices for the evaluation of the goodness of fit of SEM are based on the transformations of the difference between the estimated and the observed variance and covariance matrices. In this study the Comparative Fit Index (CFI, Bentler 1989), the Nonnormed Fit Index Tucker Lewis Index (TLI, Tucker and Lewis, 1973) and the Rooted Mean Square Error of Approximation (RMSEA, Steiger and Lind, 1980) will be considered. SEM with CFI and TLI higher than 0.95 and RMSEA lower than 0.06 are to be considered acceptable (Bollen and Long, 1993).

4. RESULTS

This section contains the results of the application of the methodologies described previously to the food quality data. The analyses were performed using *Mplus* 7 (Muthén and Muthén, 2012) to implement the SEM approach, *Winsteps* 3.75 (Linacre, 2012) for the Rasch analysis and the *R* packages *difR* (Magis *et al.*, 2010) to carry out the DIF analysis, and *ggpubr* (Kassambara, 2018) to compare the distribution of overall food quality knowledge according to students characteristics (school, attended class and food education).

4.1 *Analysis of the six scales separately*

We applied the RM and the DIF analysis for validating the six scales. For each dimension/scale, the analysis comprised the following steps: to identify problematic items, to verify the unidimensionality and local independence assumptions, to verify the reliability of the test and to examine the presence of DIF with respect to some qualitative variables describing interesting characteristics of the subjects. Reference values for the Infit and Outfit MNSQ and the first eigenvalue of PCA on standardized residuals λ_1 are reported in Wright and Linacre (1994), Smith (2002) and Brentari and Golia (2007).

Due to misfitting problems, the items Q11 of *FSec* (Infit MNSQ = 1.20, Outfit MNSQ = 1.73) and Q27 of *Nutr* (Infit MNSQ = 1.21, Outfit MNSQ = 2.17) were delated.

Table 3 reports the number of items used in the analysis, λ_1, the GLB index and the mean degree of knowledge of each dimension, with standard deviations in brackets.

TABLE 3. - *Number of items considered in the Rasch analysis, λ_1, the GLB index and the mean degree of knowledge with standard deviations in brackets*

Dimension	*# of items*	λ_1	*GLB*	*Mean (sd)*
FSaf	10	1.3	0.71	-0.21 (1.01)
FSec	9	1.5	0.78	0.45 (1.22)
Nutr	9	1.3	0.76	0.82 (1.22)
OrgQ	10	1.4	0.77	0.62 (1.10)
Sust	10	1.4	0.69	-0.12 (1.04)
SocR	10	1.4	0.91	1.04 (1.42)

The values of λ_1 are coherent with what expected in the presence of unidimensionality, and the values of GLB index are overall acceptable, meaning that the tests have an acceptable reliability. Moreover, the largest standardized residual correlations between couples of items of all the six scales range between -0.14 and -0.25, supporting the absence of local dependence.

In order to conduct the DIF analysis, we took into account three variables that stratify the sample of students in subgroups: the attended class (three levels), the attended school (six schools), which corresponds to a geographical allocation of the students, and previous food education experiences (yes or no). The test used was the Mantel-Haenszel test with its generalization in the case of multiple focal groups, the significant level was 0.05 and the adjustment for multiple comparisons was the Benjamini-Hochberg false discovery rate. Moreover, the test was performed with and without item purification, obtaining the same results. Table 4 reports the *p*-values of the DIF test for those items for which the null hypothesis of absence of DIF was rejected, with respect to the stratification variables. Only the attended class, which is related to the age, seems to induce a DIF behaviour on few items belonging to three of the six dimensions of food quality. Attended school and previous food education experiences seem not produce a DIF behaviour on any of the 58 items considered.

TABLE 4. - *Significant p-values of the Mantel-Haenszel test*

Dimension	*Variable*	*Item*	*p-value*
FSaf	Class	Q07	0.0302
FSec	Class	Q13	0.0027
	Class	Q19	0.0335
Sust	Class	Q42	0.0008
	Class	Q45	0.0138

TABLE 5. - *Measure of the items' difficulty in descending order*

FSaf		*FSec*		*Nutr*	
Item	*Measure*	*Item*	*Measure*	*Item*	*Measure*
Q8	1.96	Q18	1.36	Q24	1.24
Q7	1.76	Q19	0.83	Q30	1.02
Q9	0.92	Q13	0.68	Q21	0.41
Q3	0.80	Q14	0.50	Q26	0.37
Q5	0.19	Q12	0.10	Q28	0.33
Q10	-0.06	Q20	-0.72	Q23	-0.02
Q6	-0.74	Q16	-0.78	Q25	-0.51
Q2	-1.11	Q15	-0.80	Q22	-1.14
Q4	-1.47	Q17	-1.17	Q29	-1.70
Q1	-2.25				
OrgQ		*Sust*		*SocR*	
Item	*Measure*	*Item*	*Measure*	*Item*	*Measure*
Q36	0.73	Q43	0.82	Q59	1.03
Q40	0.68	Q48	0.46	Q57	0.87
Q35	0.59	Q46	0.44	Q56	0.81
Q34	0.53	Q44	0.33	Q55	0.60
Q38	0.17	Q50	0.30	Q53	0.12
Q31	-0.01	Q45	0.25	Q51	-0.22
Q37	-0.13	Q41	0.06	Q52	-0.27
Q33	-0.48	Q49	-0.01	Q58	-0.38
Q39	-0.95	Q42	-0.26	Q60	-1.25
Q32	-1.14	Q47	-2.41	Q54	-1.31

The performed analyses gave us indications that the scales related to *Nutr* (without item Q27), *OrgQ* and *SocR* seem valid. The average level of knowledge of these three dimensions of the food quality is higher than zero (see Table 3), suggesting that the students, overall, did not find great difficulty in correctly responding to the items. Regarding *Fsaf, FSec* and *Sust*, we found few items with DIF connected only to the attended class. Now, the presence of DIF for an item does not always indicate bias in that item; there needs to be a solid explanation for the DIF to justify the conclusion of bias (American Educational Research Association *et al.*, 2014). Given that the attended class is relevant characteristic of the children, these results suggest that these items could be revised before using the scale again.

Table 5 reports the measure of the items' difficulties in descending order. Looking at these measures, it is possible to see which facets of the six dimensions of food quality are more difficult and these findings can drive the definition of programs of food education.

Respondents expressed difficulties in matters involving the methods of storing, preparing and cooking food, and given the age of the students, it is highly plausible that many of them do not have any direct experience in food preparation but only in food consumption.

4.2 *The Overall Knowledge of food quality: validation with SEM*

To test if the proposed questionnaire can be used to measure of *OverK*, we performed a CFA based on a second-order reflective SEM; the conceptual model is shown in Figure 1a).

Each fist-order latent variable refers to one dimension of food quality and has ten reflective indicators. The second-order latent variable *OveK* has six indicators, represented by the first-order latent variables.

We first tested the conceptual model for internal consistency and unidimensionality of first-order latent variables. We deleted six indicators: Q02, Q03, Q09 related to *FSaf*, Q11 related to *FSec*, Q43, Q45 related to *Sust*. These indicators are highlighted in grey in Figure 1b). In particular for the *FSaf* dimension a revision of the three deleted indicators could be considered, firstly to verify if there is a difficulty for the children to understand the items, subsequently to find a more suitable reformulation of the facets more consistent with the related dimension.

The final version of the SEM model is shown in Figure 1b). The fit indices of the estimated SEM highlighted a good fitting of the model to data (CFI: 0.931, TLI 0.928, RMSEA 0.018). The correlation matrix of estimated factor scores (Table 6) shows a very high degree of correlation between the latent variables, consistent with the reflexive formulation of the SEM model.

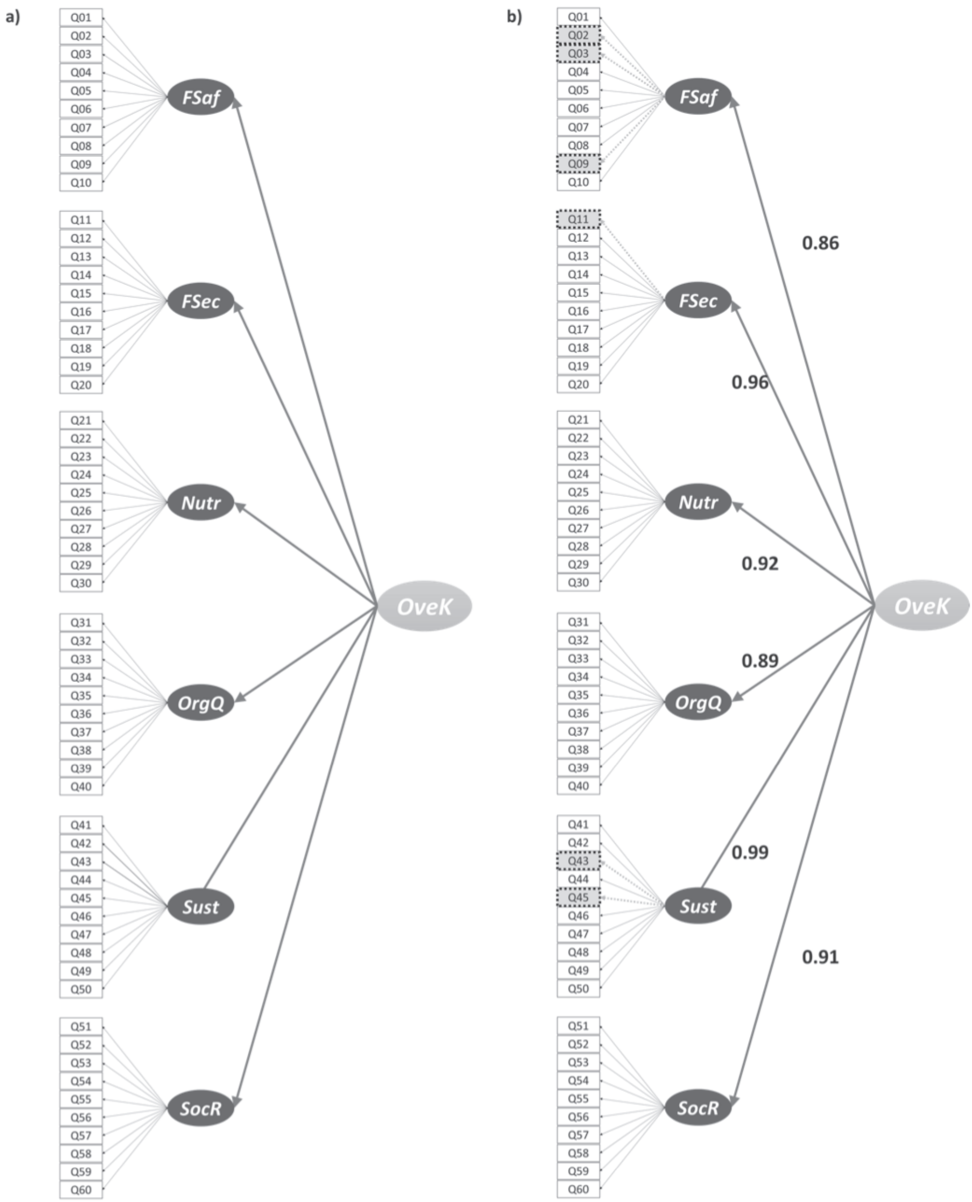

FIGURE 1. - *Conceptual framework a) and estimated SEM b). Squared boxes represent indicators, small ellipses first-order factors, large ellipse is the second-order factor. The highlighted indicators have been eliminated based on test for consistency and unidimensionality of the first-order latent variables*

TABLE 6. - *Correlation matrix of estimated factor scores*

	FSaf	*Fsec*	*Nutr*	*OrgQ*	*Sust*	*SocR*	*OverK*
FSaf	1						
FSec	0.974	1					
Nutr	0.964	0.978	1				
OrgQ	0.954	0.967	0.957	1			
Sust	0.981	0.995	0.985	0.975	1		
SocR	0.956	0.969	0.954	0.942	0.976	1	
OverK	0.982	0.996	0.986	0.976	1	0.978	1

All the results obtained by the CFA analysis confirm that the SEM conceptual model of the overall knowledge of food quality has been well specified. The fisrt-order factors of the food quality dimensions are well defined through their set of indicators (a revision of the *FSaf* dimension is suggested but it is not mandatory) and the six latent variables well represent the overall knowledge of food quality.

The scores of the *OverK* were tested to assess if the overall knowledge of food quality differs significantly among students' characteristics: school, class and previous food education experiences. Table 7 reports the results of the ANOVA analysis.

TABLE 7. - *Results of ANOVA analysis for the overall food quality knowledge*

Factor	*Df*	*Sum of squares*	*Mean of squares*	*F value*	*P-value*	
School	*5*	0.30	0.06	1.42	0.22	
Class	*2*	2.09	1.04	24.35	<0.001	***
Food education experiences	*1*	0.38	0.38	8.78	0.004	**
Residuals	*658*	28.22	0.04			

*** p-value < 0.001, ** p-value < 0.05.

The means of the groups according to the schools were not significantly different from each other, while the mean differences between the groups according to the class attended and previous food education experiences were significantly different from zero.

The factors resulted significant in the ANOVA analysis, were further analysed to assess the distribution of the overall knowledge scores within the factors (see Figure 2). The general knowledge of food quality increased with the increase of the class attended. The overall knowledge is higher for students that attended previous course on food education, pointing out that food education is relevant for the acquisition of knowledge on food quality dimensions.

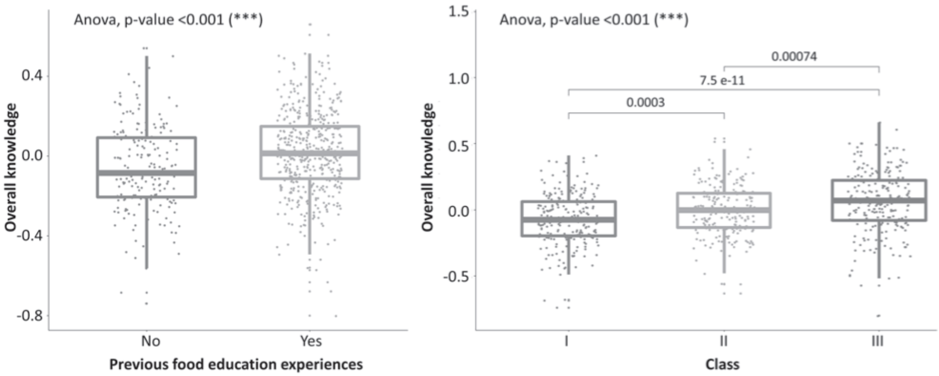

FIGURE 2. - *Results of the Wilcoxon test analysis for the pairwise comparison within the factor "Class grades" (right graph) and the factor "Previous food educational experiences" (left graph)*

Findings of ANOVA analysis underline that the knowledge of food quality is a multidisciplinary skill composed by a series of knowledge acquired transversely by students and strongly influenced by their own food experience.

5. CONCLUSIONS

Childhood is the most relevant period for the formation of the orientations in food choices. The evaluation of the actual knowledge of food quality in children is of interest for political institutions, educational agencies and health professionals, and it offers the opportunity to plan educational actions aimed at promoting the health and quality of life of today's and tomorrow's citizens. Therefore, the aim of this paper is to study how we can measure the level of knowledge of food quality in children. We chose students from lower secondary school as the target population, because the years of the lower secondary school are highly sensitive for the definition of food preferences. We followed a multidimensional approach to food quality, focusing on the following six dimensions: Food Safety, Food Security, Nutritional Value, Organoleptic Quality, Environmental Sustainability and Social Responsibility. As far as we know, there are not validated scales that can be applied, therefore we developed a scale of ten items, based on the available literature, for each of the six dimensions of food quality.

The first step of the analysis consisted in investigating the internal construct validity and reliability of the six scales, considered separately. In order to do this, we applied the RM to address several aspects related to scale validation and a DIF analysis. After the removal of two items due to misfitting problems, we found the resulting scales were unidimensional and with an acceptable reliability. The DIF analysis performed with respect to three variables, attended school, attended class and previous food education experiences, has shown that only the attended class seems to induce a DIF behaviour on few items belonging to three of the six dimensions of food quality. Given that this variable is a relevant characteristic of the children, this result suggests that the DIF items could be revised before using the scale again.

The second step of the analysis consisted in validating the overall knowledge of food quality through a multidimensional approach. To test the reliability and validity of the overall scale we performed a CFA analysis based on a Second-order SEM. The results of the analysis confirmed the defined conceptual model; therefore, it is possible to use the six scales of the questionnaire jointly to evaluate the overall knowledge of food quality. An in-depth analysis of the indicators referring to food safety could be useful although the result of the SEM model does not highlight problems on the *FSaf* dimension.

Then, we used the overall measure of the level of knowledge of food quality to segment the students with respect to attended school, attended class and previous food education experiences. We found that the attended school does not constitute a discriminant in the level of knowledge, unlike the other two characteristics. The knowledge of food quality increases with the increase of the class attended. This result underlines that food quality knowledge is a multidisciplinary skill; it consists of a series of information acquired transversely by students and strongly influenced by their own food experience. The importance of food education can be noted considering the different level of overall knowledge between students that attended previous course on food education and students that never had previous experience on it.

Based on the results of our study, we deem that the survey and the modelling techniques can be useful to design and implement food education strategies.

ACKNOWLEDGEMENT

The authors wish to thank the two anonymous reviewers for their constructive and valuable comments. This work was partially supported by "Fondazione Cariplo" (Italy) and "Regione Lombardia" (Italy) under the project: "La salute della persona: lo sviluppo e la valorizzazione della conoscenza per la prevenzione, la diagnosi precoce e le terapie personalizzate". Grant Emblematici Maggiori 2015-1080.

APPENDIX

TABLE A1. – *The fourth section of the questionnaire: percentages of correct, wrong or ''Don't know'' answer to the single items*

	Questions	**% correct answer**	**% wrong answer**	**% Don't know**
Food Safety	**Q1** - After cutting raw meat or fish it is important to wash the knife before using it again	84	4	12
	Q2 - It is important to cook eggs before eating them	68	18	14
	Q3 - Long-life milk (UHT), once opened, can be kept longer than fresh milk	30	26	44
	Q4 - In the refrigerator it is important to store food in closed containers	74	14	12
	Q5 - Cooked food can be kept well even at room temperature	42	22	36
	Q6 - It is not a problem if farm animals have access to vegetable gardens or growing fields	60	14	26
	Q7 - Washing fruits or vegetables remove all pesticide residues	16	74	10
	Q8 - In order to ensure cooked meat to be safe to eat, it is sufficient it reaches a temperature of 40º C	14	12	74
	Q9 - Thawed food cannot be refrozen if it has not been cooked	28	39	33
	Q10 - Cooked food and raw food can be stored together without problems	47	20	33
Food Security	**Q11**- Food availability in a country must always be guaranteed by the agricultural production of the country itself	17	49	34
	Q12 - Food availability is not influenced by land quality and soil degradation	56	13	31
	Q13 - The availability of food depends on the possibility of storing large quantities of food for a long time (for example, warehouses, cold stores, etc.)	45	16	39
	Q14 - Enhancing local food production contributes to food availability	48	10	42
	Q15 - All countries of the world are able to guarantee the availability of food to their citizen without aid from other countries	72	12	16
	Q16 - The possibility of obtaining the necessary food depends on the economic availability of the families	72	14	14
	Q17 - Climatic conditions are not important in determining food availability	77	13	10
	Q18 - World hunger increases as the price of cereals increases	33	27	40
	Q19 - The possibility of obtaining sufficient food for one's family depends on the proximity of market and the possibility of reaching them (eg. presence of roads, means of transport, etc.)	42	34	24
	Q20 - It is important to buy only the amount of food that will actually be consumed by the family	71	15	14

(*Continue*)

	Questions	% correct answer	% wrong answer	% Don't know
Nutritional Value	**Q21** - It is preferable, on the same day, to eat fruit and vegetables of different colors	58	18	24
	Q22 - Eating very salty foods does not cause health problems	82	10	8
	Q23 - It is preferable to use extra virgin olive oil as a condiment rather than butter	65	7	28
	Q24 - Steam cooking is the cooking method that most respects the nutritional qualities of food	42	10	48
	Q25 - Cheeses and meat are both important sources of proteins	73	12	15
	Q26 - Red meat (e.g. beef burgers) generally contains less fat than white meat (e.g. breast of chicken)	58	12	30
	Q27 - Frozen vegetables contain less nutrients than fresh vegetables	12	60	28
	Q28 - It is recommended to consume products made with whole meal flour (e.g. whole meal bread) rather than products made with refined flour (e.g. white bread)	59	12	29
	Q29 - It is important to limit the intake of fizzy drinks that are high in sugars	87	7	6
	Q30 - In a healthy diet must avoid any source of fats	46	38	16
Organoleptic characteristics	**Q31** - The method of conservation of a food does not influence its flavor	62	14	24
	Q32 - The color of a fruit can indicates the degree of ripeness	81	7	12
	Q33 - Touching a sandwich you can recognize whether it is fresh or not	71	11	18
	Q34 - Cooking vegetables change their smell and color	52	23	25
	Q35 - The unpleasant smell of a food generally affects its taste	50	32	18
	Q36 - The way food is presented influences the perception of its taste	47	23	30
	Q37 - The bitter taste of a food indicates that it should not be consumed	65	11	24
	Q38 - When you have a cold and/ or your nose is stuffed up, you perceive the taste of food differently	58	25	17
	Q39 - In choosing what you eat we also use other senses in addition to taste (sight, smell, touch, hearing)	78	12	10
	Q40 - Observing the color of the meat does not help to determine its freshness	47	20	33

(*Continue*)

	Questions	% correct answer	% wrong answer	% Don't know
Environmental Sustainability	**Q41** - Agriculture is the human activity that consumes the freshest water worldwide	46	13	41
	Q42 - The use of pesticides in agriculture hasn't any environmental impact outside the field in which they are used	52	16	32
	Q43 - The use of manure as fertilizer is controlled because it can be a source of pollution	31	36	33
	Q44 - The water needed to produce one kilogram of wheat is the same quantity needs to produce one kilogram of meat	41	9	50
	Q45 - The varieties of fruit and vegetables grown were much more numerous in the past	43	22	35
	Q46 - Pollution linked to animal breeding cannot be reduced by improving farming techniques	38	27	35
	Q47 - It is preferable to consume fruit according to the ripening season (e.g. strawberries in spring, oranges in winter, etc.)	86	5	9
	Q48 - Some types of fishing can produce important environmental impacts	38	15	47
	Q49 - It is possible to reduce the environmental impact of food production by changing one's eating habits	48	15	37
	Q50 - Through the improvement of plant cultivation techniques it is possible to reduce the impact of food production on the environment	41	13	46
Social Responsibility	**Q51** - Agricultural activities should also ensure respect for and management of environmental quality	72	6	22
	Q52 - The pay of an agricultural worker must be adequate to the work done and such as to enable him to lead a dignified life	72	9	19
	Q53 - Animal welfare in farms is not an important goal and has no influence on production	66	13	21
	Q54 - The label of a food must contain all the information useful to the consumer to know how that food is produced	85	5	10
	Q55 - Maintaining and exploiting (valorizing) local products is important as an aspect of the culture of a territory	58	10	32
	Q56 - Companies must pay attention to local habits and culture	55	16	29
	Q57 - It is not the company's responsibility to favor the education and training of its employees	54	17	29
	Q58 - Companies must guarantee their workers the opportunity to dedicate time to family and leisure	74	9	17
	Q59 - In case of need, a farm can employ workers even without a regular contract (e.g. during the harvest)	50	16	34
	Q60 - A company must guarantee its employees clean workplaces where their safety is guaranteed	84	4	12

REFERENCES

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association, Washington, DC.

Bond T.G., Fox C.M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (3rd ed.). Routledge.

Bentler P.M. (1989). *EQS structural equations program manual*. Los Angeles, CA, BMDP Statistical Software.

Bollen K.A. (1989). *Structural Equations with Latent Variables*. John Wiley and Sons, Inc., New York.

Bollen K.A., Long J.S. (1993). *Testing structural equation models* (Vol. 154). Sage.

Brentari E., Golia S. (2007). Unidimensionality in the Rasch model: how to detect and interpret. *Statistica*, **6**, 253-261.

Byrne B.M. (2013). *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming*. Routledge.

Cardello A.V. (1995). Food Quality: Relativity, context and consumer expectations. *Food Quality and Preference*, **6**, 163-170.

European Commision (2001). *Green Paper: Promoting a European framework for corporate social responsibility*. Brussels.

FAO (1996). *Declaration on world food security*. World Food Summit, Rome.

FAO (2011). *The state of the world's land and water resources for food and agriculture (SO-LAW) – Managing systems at risk*. Food and Agriculture Organization of the United Nations, Rome and Earthscan, London.

FAO (2017). *Water for sustainable food and agriculture*. A report produced for the G20 presidency of Germany. Rome: FAO.

FAO (2018). *Sustainable agriculture for biodiversity. Biodiversity for sustainable agriculture*. Rome: FAO.

FAO, WHO (2019). *Sustainable Healthy Diets: Guiding Principles*. Rome: FAO.

Flodmark C., Marcus C., Britton M. (2006). Interventions to prevent obesity in children and adolescents: a systematic literature review. *International Journal of obesity*, **30**, 579-589.

Forde C. (2018). From perception to ingestion; the role of sensory properties in energy selection, eating behaviour and food intake. *Food Quality and Preference*, 171-177.

Golia S., Simonetto A. (2015) Treating Ordinal Data: A Comparison between Rating Scale and Structural Equation Models. *Quality & Quantity*, **49**(3), 903-15.

Gustafson D., Gutman A., Leet W., Drewnowski A., Fanzo J., Ingram J. (2016) Seven Food System Metrics of Sustainable Nutrition Security. *Sustainability*, **8**, 196.

HLPE (2017) *Nutrition and Food Systems. 12. A Report by the High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security*, Rome.

Holland P.W., Thayer D.T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer and H.I. Braun (Ed.), *Test validity*, (pp. 129-145). Lawrence Erlbaum Associates.

Holsten J., Deatrick J., Kumanyika S., Pinto-Martin J., Compher C. (2012). Children's food choice process in the home environment. A qualitative descriptive study. *Appetite*, **58**(1), 64-73.

Kaplan, D. (2000). *Structural equation modeling: foundation and extensions*. Thousand Oaks: SAGE publications.

Kassambara A. (2018). *ggpubr: 'ggplot2' Based Publication Ready Plots*. R package version 0.2. https://CRAN.R-project.org/package=ggpubr.

Kim J., Oshima T.C. (2013). Effect of multiple testing adjustment in differential item functioning detection. *Educational and Psychological Measurement*, **73**, 458-470.

ISO (2014). *Discovering ISO 26000*, Genéve: International Organization for Standardization.

Linacre J.M. (2012). Winsteps® (Version 3.75.0) [Computer Software]. Beaverton, Oregon: Winsteps.com.

Magis D., Beland S., Tuerlinckx F., De Boeck P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods*, **42**, 847-862.

Mantel N., Haenszel W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, **22**, 719-748.

Marzi V. (2008) *Più ortaggi, legumi e frutta. Educazione alimentare e del gusto nella scuola primaria*. BetMultimedia.

McNeish D. (2018). Thanks Coefficient Alpha, we'll take it from here. *Psychological Methods*, **23**(3), 412-433.

Muthén B. (1983). Latent variable structural equation modelling with categorical data. *Journal of Econometrics,* **22**, 43-65.

Muthén B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, **49**, 115-132.

Muthén, L.K., Muthén, B.O. (2012). *Mplus User's Guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Penfield R.D. (2001). Assessing differential item functioning among multiple groups: A comparison of three Mantel–Haenszel procedures. *Applied Measurement in Education*, **14**, 235-259.

Penfield R.D., Camilli G. (2007). Differential item functioning and item bias. In C.R. Rao and S. Sinharay (Eds.), *Handbook of statistics: Vol. 26. Psychometrics* (pp. 125-167). Elsevier.

Perez F. (2009). *Mangiare e crescere. Aspetti cognitivi, sociali e relazionali*. Magi Edizioni.

Piqueras-Fiszman B., Alcaide J., Roura E., Spence C. (2012). Is it the plate or is it the food?

Assessing the influence of the color (black or white) and shape of the plate on the perception of the food placed on it. *Food Quality and Preference*, **24**(1), 205-208.

Poore J., Nemecek T. (2018). Reducing food's environmental impacts through producers and consumers. *Science*, **360**(6392), 987-992.

Popkin B.M., Corvalan C., Grummer-Strawn L.M. (2020). Dynamics of the double burden of malnutrition and the changing nutrition reality. *The Lancet*, **2020**, 65-74.

Rasch G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen: Danish Institute for Educational Research (Expanded edition, Chicago: The University of Chicago Press, 1980).

Simonetto A. (2012). Formative and reflective models: state of the art. *Electronic Journal of Applied Statistical Analysis*, **5**(3), 452-457.

Smith, Jr.E.V. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, **3**, 205-231.

Steiger J.H., Lind J.C. (1980). Statistically based tests for the number of common factors. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Somes G.W. (1986). The generalized Mantel-Haenszel statistic. *The American Statistician*, **40**, 106-108.

Tucker L.R., Lewis C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, **38**, 1-10.

Vågsholm I., Arzoomand N.S., Boqvist S. (2020) Food Security, Safety, and Sustainability - Getting the Trade-Offs Right. *Frontiers in Sustainable Food Systems* 4.

WHO (2006). *Five keys to safer food*, France: WHO.

WHO (2015). *Health 2020: Agriculture and Health through Food Safety and Nutrition*, available at: http://www.euro.who.int/__data/assets/pdf_file/0016/324610/Health-2020-Agriculture-and-health-through-food-safety-and-nutrition-en.pdf?ua=1, last access: 18 May 2020.

WHO (2018). *Guideline: implementing effective actions for improving adolescent nutrition*. WHO, Geneva.

WHO, FAO (2012). *Five keys to growing safer fruits and vegetables: promoting health by decreasing microbial contamination*, Switzerland: WHO; FAO.

Willett W., Rockström J., Loken B., Springmann M., Lang T., Vermeulen S., Jonell M. (2019). Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems. *The Lancet*, **393**(10170), 447-492.

Wright B.D., Linacre J.M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, **8**(3), 370.

Wright B.D., Master G.N. (1982). Rating scale analysis, MESA Press, Chicago Bollen, K.A. (1989). *Structural Equations with Latent Variables*. John Wiley and Sons, Inc., New York.