Reliability generalization meta-analysis of the food neophobia scale: Turkish sample

Eray Polat
Gumushane University,Turkey

Serkan Çalışkan
Kastamonu University,Turkey

Abstract:

***Purpose**: The Food Neophobia Scale developed by Pliner and Hobden (1992) has been widely utilized globally and in Turkey for many years to measure people's fear of new foods. This study aims to determine the mean value by conducting a reliability generalization meta-analysis for the reported reliability coefficients of individual studies in Turkey's tourism field, which employed Pliner and Hobden's scale to investigate food neophobia. Additionally, this study explores variations in the mean value among subgroups.*
***Methods**: A reliability generalization meta-analysis based on a random-effects model was conducted to examine the heterogeneity of reliability coefficients in the study, along with heterogeneity analyses and moderator analyses.*
***Results**: Based on the analysis of 48 independent samples (N= 23306), the transformed mean Cronbach's alpha value was estimated to be .827 (95% CI [.796-.853]) and found to be significant. The Q-test and I2 values reveal significant heterogeneity between alpha coefficients, indicating a notable variation in the measurement reliability across samples. Moderator analyses using analog to the ANOVA and meta-regression analyses showed that reliability coefficients differed according to the variables of publication type, sample type, and proportion of women in the sample.*
***Implications**: The results offer valuable insight for researchers seeking to select appropriate scales for investigating food neophobia.*

Keywords: *food neophobia, reliability generalization, meta-analysis*

JEL Classification: *L66, L83, Z33*

Biographical note: Eray Polat is an Associate Professor at the Faculty of Tourism at Gumushane University in Turkey. He holds a PhD degree in tourism management from Balikesir University, Turkey. His research interests focus on job insecurity, ethics, and systematic reviews in tourism and hospitality. Serkan Çalışkan (serkancaliskan06@gmail.com) is an Assistant Professor at the Faculty of Tourism, Kastamonu University. He holds a PhD degree in tourism management from Balikesir University, Turkey. His areas of interest are hospitality management, gastronomy tourism, gastronomy education.

1 INTRODUCTION

Food, a necessary component for the survival of all living organisms, has been studied by numerous academic disciplines (Rabadán & Bernabéu, 2022). Nutrition, the fundamental requirement for all living organisms, strives to safeguard and enhance well-being by deliberately consuming essential nutrients. Eating habits are formed through the repetition of nutritional behavior, which occurs multiple times daily. The resulting repetition of these activities leads to the development of habits, providing an individual with a sense of security and comfort from the familiar, perceived as low risk. Pliner et al. (1993) reported that participants rated familiar foods as more palatable, which influenced their motivation to taste the food. Conversely, food neophobia (FN), which refers to individuals' reluctance or fear to try new foods, involves attitudes and behaviors toward tasting or avoiding novel foods. Fallon and Rozin's (1983) classification form the underpinnings of this phenomenon. This classification focuses on the potential hazards related to negative emotions such as dislike, disgust, and fear. Eating and drinking preferences serve as key elements of daily life, functioning as a marker of an individual's identity. It is therefore crucial for marketers, food producers, and all tourism industry stakeholders to comprehend why certain foods are readily accepted or rejected.

Although FN has been widely studied, the specific mechanisms for it remain unclear (Lafraire et al., 2016). Accurately measuring the various aspects of FN and willingness to try new foods requires using appropriate scales (Rabadán & Bernabéu, 2022). Various scales have been developed to assess these different aspects, highlighting the need for standardized methods of measurement. Damsbo-

ISSN: 2529-1947. UDC: 658.8+338.48+339.1+640(05)
http://doi.org/10.5281/zenodo.13772035
Published online: 30 June 2024
www.jthsm.gr

Svendsen and colleagues (2017) investigated 13 scales that measured FN through participants' interest in trying novel foods. The scales consist of 6 to 35 items, with Cronbach's alpha values ranging from 0.80 to 0.92. Participant responses were collected using a 3-7 Likert rating. In scale development studies, the sample size has been reported to range from 133 (Pliner et al., 1993) to 16,644 (Kaiser et al., 2012) and includes individuals from various age groups, including children (Rubio et al., 2008), young adults (Raudenbush et al., 1995), and adults (Frank, 1994), with ages ranging from 2 to 65 years.

FN has been extensively researched for over 40 years due to its impact on food quality and variety. In 1992, Pliner and Hobden introduced an instrument to measure FN in individuals, viewing avoidance of new foods as a personality trait. The Food Neophobia Scale (FNS) is a unidimensional scale comprising 10 items. The scale yields scores ranging from 10-70, wherein high scores indicate neophobia or fear of new foods, while low scores indicate food neophilia or openness to try new foods. This scale, which was validated with a sample of undergraduate psychology students in Canada, has been applied to many different groups in the following years: children (Skinner, 2002), families (Koivisto, 1997), consumers from different geographies (Murray, 2001), obese (Monneuse et al., 2008), pregnants (Paupério et al., 2014), tourists (Akin et al., 2023; Sivrikaya, 2020) or kitchen staff in hotels (Üngüren & Tekin, 2022).

The FNS is a commonly used tool to measure FN across diverse groups. However, worldwide application of the FNS, which was developed and validated on specific and dissimilar samples, poses various challenges (Ritchey et al., 2003). To address these issues, some researchers have utilized techniques such as modifying certain terms (Elkins & Zickgraf, 2018), removing items from the scale (Sogari et al., 2019), or adjusting the Likert scale utilized (Dönmez & Sevim, 2023).Since its introduction in 1992, FNS has been used in a wide variety of studies. For example, studies assessing the impact of dietary diversity on human health (Costa et al., 2019, Jaeger et al., 2017); studies focusing on new product development or reformulation (Domínguez et al., 2019; Rabadán et al., 2021); sensory analysis studies (Chung et al., 2012, Reverdy et al., 2008); studies examining the impact of socioeconomic characteristics of individual FNS (Meiselman et al., 2010; Nordin et al., 2004). Pliner & Hobden (1992) determined the scale's Cronbach's alpha to be 0.88 in the original study. Jaeger et al. (2017) reported a Cronbach's alpha coefficient of 0.83 in their adult study of 1167 participants in New Zealand. In Italy, Laureati et al. (2018) found a coefficient of 0.87 with 1225 participants. For the overall scale, Siegrist et al. (2013) found coefficients of 0.80 with a sample of 4442 participants, while German and French speakers had coefficients of 0.79 and 0.82, respectively. Bernal-Gil et al. (2020) found Cronbach's alpha to be 0.612 in a study that examined Mexican consumers' (n=160) opinions on ethnic cuisine.

The FNS's Turkish validity and reliability study was conducted with a sample of 444 participants by Uçar et al. (2021). The study resulted in a two-dimensional structure with a Cronbach's alpha coefficient of 0.805. In another study where the scale was adapted into Turkish, Duman et al. (2020) used data from 195 adult patients and found the Cronbach's alpha coefficient to be 0.614. In both studies, it was determined that the FNS is valid and reliable in a Turkish sample. Other studies conducted in the Turkish literature reported varied results. Specifically, Konaklıoğlu and Algül (2022) found a Cronbach's alpha coefficient of 0.6 in their study with a sample size of 585 participants, while the highest coefficient was 0.971 in Öztürk's (2019) research.

Previous studies have shown inconsistent reliability coefficients for the FNS. However, there has been no research that explores the generalizability and heterogeneity of the FNS reliability coefficients. It is necessary to investigate this in order to estimate the FNS's overall reliability and to understand the reasons behind the heterogeneity in reliability coefficients across different samples (Sen, 2022). A possible approach for this examination is conducting a meta-analysis of reliability generalization. Meta-analysis is a quantitative method that synthesizes results from previous studies on a specific research topic to draw a single conclusion. In essence, meta-analysis combines findings from multiple studies and presents them as a single result (Aslan et al., 2022). Additionally, meta-analysis studies provide the opportunity to explain possible inconsistent findings in previous research and the moderators that may have caused them, leading to a more harmonious expression of results (Çelik, 2023; Çelik et al., 2023). Meta-analysis studies focused on the reliability values of a specific scale are referred to as reliability generalization meta-analyses (Vacha-Haase, 1998). This meta-analysis examines the variation in reliability values of test measurements across diverse sample sets and levels of heterogeneity (Badenes-Ribera et al., 2023; Sanchez-Meca et al., 2021).

One of the most commonly used scales in FN studies is the FNS (Rabadán & Bernabéu, 2022). The primary aim of this study is to establish the average reliability value of the FNS, designed by Pliner and Hobden (1992), which is extensively applied in the literature for the Turkish sample based on tourism studies. Additionally, this study examines how this value may fluctuate among distinct subgroups through moderator analyses. In this direction, the present study will answer two primary research questions: 1) What is the overall reliability coefficient of the FNS? and 2) Are there any differences in reliability scores among subgroups?

The research will be particularly useful for researchers. Gastronomy tourism and food-related components of the tourist experience are promoted and sought after as distinctive aspects worldwide. There will often be visitors who are hesitant or afraid to experience the food in the destination. Therefore, there will continue to be significant interest in food neophobia research going forward. The selection of scales for these studies is vital for the research's health and quality. In this regard, our findings on FNS will offer significant perspectives for gastronomy and FN researchers.

2 METHODOLOGY

The study aims to establish the reliability generalization of the FNS. The steps outlined in the REGEMA flow diagram (Sanchez-Meca et al., 2021) were followed to guide the writing of the reliability generalization meta-analysis.

2.1. Selection criteria

The dataset used in this study consists of previous research conducted in Turkey that measured FN using the scale developed by Pliner and Hobden in 1992. Table 1 presents the criteria utilized during this process. The timeframe under consideration spans from 1992 to 2023 (October 18, 2023) due to the initial publication of the scale in 1992. This study includes relevant scientific articles and theses written in English and Turkish that focus on the tourism discipline in Turkey. Finally, the meta-analysis only excluded reliability coefficients other than Cronbach's alpha, including omega and test-retest.

Table 1. Selection

Criterion	Inclusion
Time invertal	All publications between 1992-2023
Publication type	Scientific articles and theses
Language	Turkish and English
Compliance	Quantitative studies using the FNS and conducted in Turkey
Discipline	Tourism
Reliability type	Cronbach's alpha

2.2. Search strategies and data extraction

National Thesis Center of Turkey (YOKTEZ), DergiPark, and Google Scholar were utilized to access studies within the research scope. The search term "food neophobia" was used to gather studies. As all the studies had English abstracts, studies conducted in Turkey were also accessible. As of October 18th, 2023, a total of 241 Turkish studies [Google Scholar (n= 176), DergiPark (n= 23), YOKTEZ (n= 42)] were examined.

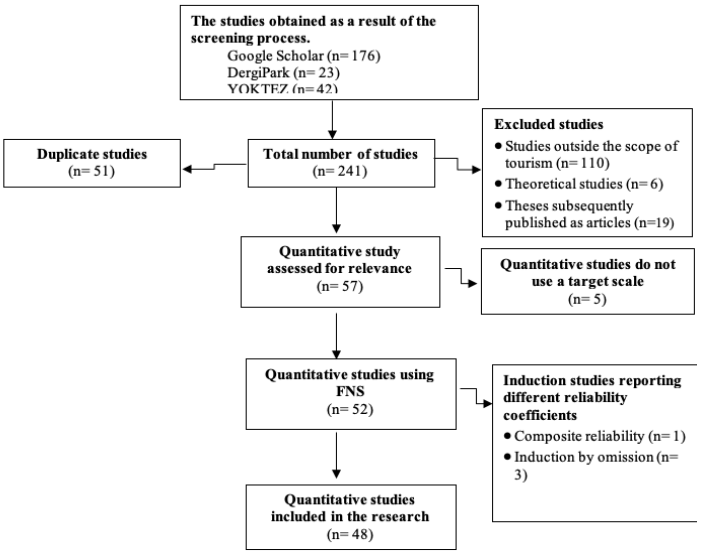

Figure 1. REGEMA flow diagram

As shown in Figure 1 of the REGEMA diagram, the literature review resulted in the elimination of a total of 184 studies. Of these, 51 appeared in the searched databases simultaneously and were removed due to duplicate publications. Additionally, 108 studies fell outside the realm of tourism (e.g., health, food engineering, etc.), while six studies were purely theoretical and did not report alpha values. Finally, 19 studies were identified as theses that were subsequently published as articles. Out of the 57 studies that remained, five did not utilize the relevant scale, one used a coefficient other than Cronbach's alpha as a measure of reliability, and three were eliminated because they did reliability induction by omission (Sanchez-Meca et al., 2021). The researchers who did reliability induction by omission were contacted by e-mail, but no response was received. Subsequently, 48 specific studies that adhered to the research criteria were analyzed.

2.3. Procedure and coding

The study determined Cronbach's alpha values for 48 samples within the final data set. To identify possible factors influencing the change in alpha values, moderator variables were selected based on previous research (Aslan et al., 2022; Çelik, 2023; Sen, 2022; Yörük & Sen, 2022), and the data was coded using Microsoft Excel. Summary information regarding the coded variables is presented in Table 2.

To ensure inter-coder reliability, the second researcher checked the coding conducted by the first researcher of the study. Differences in coding emerged across the three studies. As in previous similar studies (Çelik et al., 2023; Polat & Koseoglu, 2022), the researchers deliberated on these discrepancies and reached a consensus, bringing the process to completion.

Table 2. Variables and coding criteria based on the coding process

Variable	Type	Coding criteria
Study ID	Categorical	The number assigned to the study
Year of publication	Continuous	Year of publication
Publication type	Categorical	Thesis, Article
Sample size	Continuous	Sample size in the study
Cronbach's alpha coefficient	Continuous	Alpha value determined in the study
Number of items	Continuous	Number of items in the scale used
Sample type	Categorical	Students, tourists, locals
Scale mean	Continuous	The average score obtained from the scale
Method of application of the scale	Categorical	Online, face to face, hybrid
Percentage of female participants	Continuous	The proportion of female participants in the sample

2.4. Statistical analyses

Since reliability coefficients in individual studies typically exhibit a skewed distribution (Semma et al., 2019), alpha coefficients require transformation in computing the average reliability value. The Bonett Transformation Formula (2002) ($T= \ln(1-|\alpha|)$) was deployed for this purpose. The Bonett transformation normalizes the skewed distribution of alpha coefficients and balances their variances. In order to interpret the obtained results, it is necessary to convert the obtained value back to alpha values (Çelik, 2023). The meta-analysis literature utilizes either the fixed-effect model or the random-effects model (REM) to calculate the average effect size (Borenstein et al., 2021). REM offers more precise confidence intervals compared to other models, making its assumptions more realistic in social sciences than the fixed-effect model (Sen, 2022).

Heterogeneity among alpha coefficients was assessed using Q statistics and I2 values. Significant Q statistics and I2 values over 75% suggest the presence of heterogeneity (Özdemir et. al, 2020). Additionally, PI values were also included to assess heterogeneity. The PI allows for future research to have an idea of the true range of effects to expect and presents heterogeneity using the same metric as the original effect size measure (IntHout et al., 2016; Morris,

2023; Nunez-Nunez et al., 2022). By analog to the ANOVA, the effect of categorical moderator variables on the heterogeneity of reliability estimates was examined. The moderator variables include publication type (thesis, article), sample type (students, tourists, locals), and the way the scale is administered (online, face-to-face, mixed). The study also employs a meta-regression approach to investigate the impacts of continuous variables such as publication year, sample size, number of items, scale mean, and gender-female ratio (%). Finally, the R2 estimates were utilized to clarify the amount of variation contributed by the moderator variables (Yörük & Sen, 2022).

Given the inclusion of only published and accessible studies in the analysis, unreported reliability coefficients may have adverse effects on the present meta-analysis results (Sen, 2022). Therefore, it is crucial to examine the potential for publication bias. To investigate this, the trim-fill method developed by Duval and Tweedie (2000), which involves creating a funnel plot, was initially used. This method estimates the potential number of missing studies in the meta-analysis and their effects on the findings (Guzeller & Celiker, 2020). The symmetrical distribution of publications in the funnel plot exhibits no publication bias. If any missing studies are found, they will be placed in the appropriate locations (to the right or left) on the graph to create a symmetrical appearance (Çelik, 2023) and to demonstrate the impact of missing studies on the results. Other methods have been employed to explore publication bias, including Rosenthal's (1979) and Rosenberg's (2005) fail-safe N, Egger's regression test, and Begg and Mazumdar's (1994) rank correlation test. All statistical analyses were conducted using the metaphor package in the R software and the MAJOR module in Jamovi

3 FINDINGS

3.1. Overview of individual studies

Although the scale, which was validated in the study, was developed in 1992, it was not utilized in studies regarding tourism in Turkey until 2016. Nonetheless, it is apparent that the attention given to the scale and the topic of food neophobia has recently heightened, particularly in 2021 (refer to Figure 2).

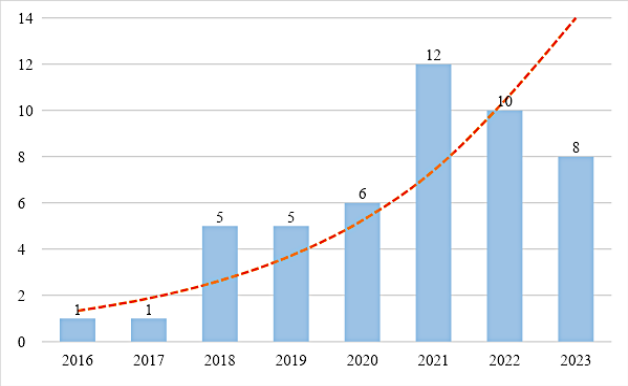

Figure 2. Distribution of individual studies by year

Table 3 displays additional characteristics of the studies. The majority of studies (62.5%) utilized a 10-item scale consistent with the original scale. Raw alpha values ranged from ,60 to ,97 (M= ,80; sd=,09). Additionally, the sample size fluctuated between 205 and 1286 (M= 458.5; sd= 197.6; median= 412.5).

Table 3. General characteristics of the studies included in the data set

	Number of studies (k)	***%***
Publication type		
Thesis	25	52,1
Article	23	47,9
Number of items		
10 items	30	62,5
<10	18	37,5
Sample type		
Tourist	24	50
Locals	16	33,3
Student	8	16,7
Method of application of the scale		
Face to face	32	66,7
Online	10	20,8
Mix	6	12,5
	Mean (M)	***Standard deviation (sd)***
Percentage of female participants	54,22	10,40
Sample size	485,5	197,6
Raw alpha value	0,80	0,09

3.2. Reliability generalization and heterogeneity

The study utilized Bonett's (2002) formula with REM to perform analyses with alpha values. Results showed a statistically significant transformed alpha value of ,827 (95% Confidence Interval (CI) [,796-,853]) with a p-value < ,001. High heterogeneity was observed in the reliability estimates as indicated by the Q statistic (Q(47) = 3735.6794; p < ,0001) and the I2 value (98.6%). In addition, the 95% PI interval was wide (.454-945), indicating significant uncertainty regarding the expected reliability in future studies using the FNS.

3.3. Assessment of publication bias

Numerous methods were employed to detect any publication bias. Initially, the funnel plot (Figure 3) was assessed, and it was found that the symmetry was somewhat distorted, indicating the possible omission of some studies on the right side. To address this issue, the trim-and-fill method developed by Duval and Tweedie (2000) was employed, which resulted in the addition of 11 studies on the right side of the graph. The newly calculated mean reliability estimate, inclusive of these additional studies, was found to be statistically significant ($p<0.001$) at 0.858 with a 95% CI of [0.831 - 0.880]. Upon comparison with the previous reliability estimate obtained prior to the implementation of the trim-and-fill method, there was a slight difference of 0.031.

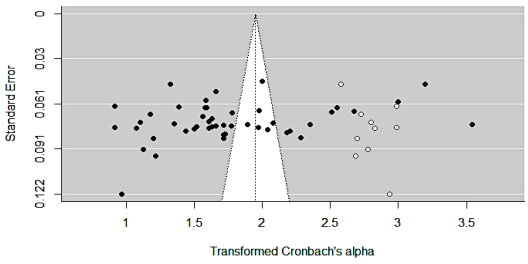

Figure 3. Funnel plot obtained by applying the trim-and-fill method

In addition, the Egger regression test yielded an insignificant t-value (t[46] = -1.5256; p=0.134). Furthermore, according to the Begg and Mazumdar (1994) rank correlation test, the Kendall's Tau value between alpha coefficient values and standard errors is both negative and insignificant (τ = -0.059; p = 0.557). These statistically insignificant p-values demonstrate the absence of publication bias. Finally, upon analyzing Rosenthal's (1979) and Rosenberg's safe N values, 545,806 and 409,328 studies are required, respectively, for the average alpha value to become statistically insignificant (p > .05). These values are more significant than 5k+10 (Sen, 2022) (=240), indicating no publication bias. When considering the results as a whole, it is evident that the research poses no risk of publication bias.

3.4. Relationship between the moderating variables and reliability estimate

In the study, moderator variable analyses were conducted to explain heterogeneity. In this analysis, alpha value is the dependent variable, categorical (type of publication, type of sample, type of application of the scale) and continuous (year of publication, sample size, number of items, scale means, gender-female ratio (%)) variables are independent variables. Continuous variables were analyzed using meta-regression analysis, while categorical variables were analyzed using an analog to the ANOVA.

Table 4. Moderator analysis results for categorical variables (Analog to the ANOVA)

Variable	Category	*k*	*N*	*α+*	*%95 CI*	ANOVA results
Publication type	Article	23	9935	,775	,736 - ,814	F (1, 46) = 4,1709, p= ,047*
	Thesis	25	13371	,829	,793 - ,866	R^2= 6,73
						Qw (46) = 2682,5865 p<,0001
Sample type	Tourist	24	12917	,807	,770 - ,844	F (2, 45) = 3,4086, p= ,042*
	Locals	16	7632	,834	,789 - ,879	R^2= 9,31
	Student	8	2757	,731	,665 - ,797	Qw (45) = 3580,7556 p<,0001
Method of application of the scale	Face to face	32	16228	,815	,754 - ,876	F (2, 45) = ,5499, p= ,58
	Online	10	3922	,765	,685 - ,845	R^2= 0
	Mix	6	3156	,807	,773 - ,841	Qw (45) = 3543,3197 p<,0001

Note: k: number of studies; N: total sample size; α+: mean alpha value; CI: confidence interval; F: Knapp-Hartung statistic testing the significance of the moderator variable; Qw: statistic testing the significance of the model; R2: proportion of variance explained by the moderator variable.

No statistically significant difference was found among the subcategories in the way the scale was applied, in terms of categorical variables (Table 4) (p>0.05). However, the reliability estimates were significantly affected by the type of publication (thesis & article) (p=0.047; R2=6.73) and the type of sample (tourist & locals & students) (p=0.042; R2=9.31). These results indicate that alpha values vary within the subcategories and account for the significant heterogeneity observed in the alpha coefficient estimates. In regard to publication type, the mean reliability estimate (,829) for thesis studies (k=25) surpasses the mean reliability estimate (,775) of article studies (k=23). The average reliability estimate calculated based on sample type presents a higher value for the local people sample (,834) compared with the tourist sample (,807), while the student sample has the lowest value (,731).

Table 5 displays the moderator analyses for continuous variables. The results indicate that year of publication (p= ,52), sample size (p= ,24), number of items (p= ,27), and scale mean (p= ,80) do not serve as significant predictors. However, the proportion of women in the sample is statistically significant at a 10% level (p= ,08). It was observed that as the proportion of women in the sample increases, the reliability value decreases by 0.0024.

Table 5. Results of Moderating Test of Brand Awareness in the Social Media Marketing and Brand Loyalty Relationship

Variable	*k*	*bj*	*SE*	*%95 CI*	*F*	*p*	*Qe*	R^2
Year of publication	48	,005	,0078	-20,744 – 42,655	,416	,52	4044,188	0
Sample size	48	,000	,000	-,0001 - ,0002	1,372	,24	3955,1826	,05
Number of items	48	,0064	,0059	-,0054 - ,0182	1,207	,27	3654,8237	,01
Scale mean	39	-,0093	,0380	-,0863 - ,0678	,0593	,80	3457.6811	0
Gender (female rate %)	47	-,0024	,0013	-,005 - ,0003	3,167	**,08****	4046,5842	4,74

Note: k: number of studies; bj= unstandardized regression coefficient; SE= standard error; CI: confidence interval; F = Knapp-Hartung statistic testing the significance of the moderator variable; Qe= heterogeneity statistic; R2 = explained variance ratio.

Finally, a weighted multiple meta-regression model was constructed to explicate the calculated heterogeneity in reliability estimates. The resulting final predictor model incorporated three moderator variables: the type of publication, the sample type, and the percentage of females in the study population. The complete model (Qe= 2578.8330; p< ,0001) and the Knapp-Hartung statistic, which tests the significance of the moderator variables (F(1, 42)= 2.705; p= ,04), were both statistically significant. The constructed model explains 13.24% of the total variance. The findings demonstrate that variables beyond the three predictors exert a significant impact on the alteration of alpha values.

4 DISCUSSION

The purpose of this reliability generalization meta-analysis is to acquire the overall reliability of the FNS created by Pliner and Hobden (1992) in the field of tourism within the Turkish sample. Furthermore, the study aims to analyze moderator variables that aid in comprehending the inconsistencies among studies. To achieve this objective, individual studies were analyzed which utilized the target scale and reported the alpha coefficient. Given the heterogeneity of alpha values in these studies (ranging from .60 to .97), it is necessary to determine an average alpha value. The combined reliability coefficient, obtained from a total of 48 independent samples, was found to be .827, which is lower than the value (.88) reported in the original study conducted by Pliner and Hobden in 1992. The inclusion of local populations and students (50%) in the analyzed studies, as well as cultural disparities based on these groups, may have played a role in this outcome.

Based on the mean value obtained in this study, the overall estimate of Cronbach's alpha falls within rational limits, surpassing the recommended threshold of >.70 for exploratory and >.80 for general research (Nunnally & Bernstein, 1994). Nonetheless, this average value is not appropriate for clinical research (> .90) (Badenes-Ribera et al., 2023). However, the scale's reliability values (PI) may range from ,454 to ,945 in future studies. This indicates that the assessed scale may not be very stable and may exhibit

significant heterogeneity. Moreover, this value could vary depending on the sample's average age, gender, or experience and may fall below the lowest reasonable threshold of .70. Thus, it may be inappropriate for both exploratory and general research as well as clinical applications. Therefore, researchers investigating FN should consider the variables identified in this study that impact heterogeneity, while utilizing the Pliner and Hobden (1992) scale. Additionally, attention should be given to the sampling methods employed in order to achieve more reliable coefficients of reliability.

High heterogeneity was observed across the studies (I2= 98.6%). Consequently, reliability coefficients obtained from the FNS scale vary depending on the samples and cannot be generalized. Therefore, researchers should not use reliability induction for this scale. Moreover, it is crucial to report reliability after all measurements as it is a result of measurements rather than a test characteristic (Bademci, 2004).

Moderator analyses revealed three factors (publication type, sample type, and proportion of women) that may account for the heterogeneity in reliability coefficients. It is important to note that the mentioned variables should be considered when interpreting reliability coefficients. Examination of publication type demonstrated higher reliability estimates in thesis studies. This may be due to researchers being more meticulous during data collection or adhering more closely to sampling rules. More specifically, thesis studies may have excluded convenience sampling groups, such as undergraduate students who lacked motivation to consistently respond to self-report questionnaires (Phillips et al., 2016). In fact, the lowest reliability coefficient was identified among the sample of students. However, despite a variance in the reliability coefficients for thesis and article research, both values remain within acceptable limits.

Interestingly, the reliability of the scale decreases as the percentage of women in the sample increases. It is worth noting that this result was observed at a reliability level of 10%. The reduction in the reliability coefficient as the proportion of women increases might suggest that the FNS is more suitable for men, or at least for populations with a significant male presence. When examining sex ratio, lower alpha estimates among women indicate a higher level of measurement error (Shou & Olney, 2020). One possible explanation is that greater disgust sensitivity in women results in higher neophobia costs (Çınar et al., 2021). This may introduce additional measurement error among women, thereby leading to a lower alpha coefficient (Shou & Olney, 2020). While there are various psychometric studies on the scale in the literature (Fernández-Ruiz et al., 2013; Ritchey et al., 2003; Zhao et al., 2020), no research has examined the measurement invariance of the scale across genders, to the best of the researchers' knowledge. Conducting such a study could offer further evidence to endorse the use of the scale in female populations in future research (Vassar, 2008).

The final multiple meta-regression model developed to account for the heterogeneity in reliability estimates explicated 13.24% of the total variance. Hence, reckoning other factors that may account for this heterogeneity is imperative. The study selected moderator variables based on the descriptive information provided in individual studies. However, the potential effects of values such as the average age of the sample or the standard deviation of the scale mean, which could act as moderators, on heterogeneity could not be examined due to their absence in the studies. Hence, future research is recommended to investigate new moderator variables and discern their influences.

4.1. Limitations

Despite its significant contributions, this study has limitations. Firstly, it should be noted that this study has only included studies conducted in Turkey. However, future studies that consider all relevant literature will allow for a more comprehensive determination of the Cronbach's alpha value and facilitate cross-cultural comparisons. Secondly, although a rigorous data collection process was implemented, there is a possibility that some studies were not included. Therefore, to replicate the study, other databases like Scopus, Web of Science, and EBSCO must be included. Furthermore, the limited representativeness of the study may be attributed to the exclusion of studies with low alpha values from publication. As a result, the Cronbach's alpha values obtained are restricted to the literature review's scope. Additionally, missing demographic data in some studies has emerged as another limitation of this research. This may have led to an inability to investigate possible moderators of heterogeneity. One potential limitation of the study could be the implementation of Bonett's transformation for standardizing alpha values. Subsequent research could aid in improving our comprehension of the pertinent variables involved in achieving reliability across studies by utilizing alternative transformation techniques (such as Hakstian-Whalen or Fisher Z) or examining other measures of reliability (such as test-retest reliability).

5 CONCLUSION

In conclusion, this study demonstrates that the FNS maintains acceptable values, on average. However, it is important for future studies utilizing this scale to remain vigilant and implement measures to prevent potential factors that may decrease reliability, such as the composition of the student sample or the proportion of women involved. Additionally, researchers utilizing the scale should consider generating their own estimates of reliability based on their individual data rather than relying on induction.

Acknowledgement: We would like to express our gratitude to Fatih Çelik, PhD, from Trabzon University, Turkey, for his invaluable assistance in the data analysis process.

REFERENCES

Akin, A., Akin, A., & Mutlu, H. T. (2023). Are tourists neophobic against offal meals? International Journal of Gastronomy and Food Science, 31, 100684. https://doi.org/10.1016/j.ijgfs.2023.100684

Aslan, Ö., Göcen, S., & Şen, S. (2022). Reliability Generalization Meta-Analysis of Mathematics Anxiety Scale for Primary School Students. Journal of Measurement and Evaluation in Education and Psychology, 13(2), 117-133.

Bademci, V. (2004). “Testin güvenirliği” veya “test güvenilirdir” diye ifade etmek doğru değildir. Türk Eğitim Bilimleri Dergisi, 2(3), 367-373.

Badenes-Ribera, L., Duro-García, C., López-Ibáñez, C., Martí-Vilar, M., & Sánchez-Meca, J. (2023). The Adult Prosocialness Behavior Scale: A reliability generalisation meta-analysis. International Journal of Behavioral Development, 47(1), 59-71. https://doi.org/10.1177/01650254221128280

Begg, C. B., & Mazumdar, M. (1994). Operating Characteristics of a Rank Correlation Test for Publication Bias. Biometrics, 50(4), 1088-1101. https://doi.org/10.2307/2533446

Bernal-Gil, N. Y., Favila-Cisneros, H. J., Zaragoza-Alonso, J., Cuffia, F., & Rojas-Rivas, E. (2020). Using projective techniques and Food Neophobia Scale to explore the perception of traditional ethnic foods in Central Mexico: A preliminary study on the beverage Sende. Journal of Sensory Studies, 35(6), e12606. https://doi.org/10.1111/joss.12606

Bonett, D. G. (2002). Sample Size Requirements for Testing and Estimating Coefficient Alpha. Journal of Educational and Behavioral Statistics, 27(4), 335-340. https://doi.org/10.3102/10769986027004335

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). Introduction to Meta-Analysis. John Wiley & Sons.

Chung, L., Chung, S.-J., Kim, J.-Y., Kim, K.-O., O'Mahony, M., Vickers, Z., Kim, H.-R. (2012). Comparing the liking for Korean style salad dressings and beverages between US and Korean consumers: Effects of sensory and non-sensory factors. Food Quality and Preference, 26(1), 105–118. https://doi.org/10.1016/j.foodqual.2012.03.011

Costa, A., Silva, C., & Oliveira, A. (2020). Food neophobia and its association with food preferences and dietary intake of adults. Nutrition & Dietetics, 77(5), 542-549. https://doi.org/10.1111/1747-0080.12587

Çelik, F. (2023). Gelişmeleri Kaçırma Korkusu Ölçeği Türkçe Versiyonunun Güvenirlik Genelleme Meta Analizi. Gümüşhane Üniversitesi İletişim Fakültesi Elektronik Dergisi, 11(2), 848-877.

Çelik, F., Çam, M. S., & Koseoglu, M. A. (2023). Ad avoidance in the digital context: A systematic literature review and research agenda. International Journal of Consumer Studies, 47(6), 2071-2105. https://doi.org/10.1111/ijcs.12882

Çınar, Ç., Karinen, A. K., & Tybur, J. M. (2021). The multidimensional nature of food neophobia. Appetite, 162, 105177. https://doi.org/10.1016/j.appet.2021.105177

Damsbo-Svendsen, M., Frøst, M. B., & Olsen, A. (2017). A review of instruments developed to measure food neophobia. Appetite, 113, 358-367. https://doi.org/10.1016/j.appet.2017.02.032

Domínguez, L., Fernández-Ruiz, V., Sánchez-Mata, M. C., & C´amara, M. (2019). Food neophobia: Spanish case study related to new formulations based on traditional 'gazpacho'. Acta Horticulturae, (1233), 209–216.

Dönmez, Ö. Ç., & Sevim, B. Yükseköğretim kurumlarının gıda ile ilgili bölümlerindeki öğrencilerin neofobi düzeylerinin belirlenmesi. GSI Journals Serie A: Advancements in Tourism Recreation and Sports Sciences, 6(2), 288-300. https://doi.org/10.53353/atrss.1229220

Duman, E., Akçil Ok, M., Keser, A. (2020). Besin neofobisi ölçeğinin Türkçeye uyarlanması: geçerlik ve güvenirlik çalışması. Kocaeli Üniversitesi Sağlık Bilimleri Dergisi, 6(2),157–61. https://doi.org/10.30934/kusbed.643038

Duval, S., & Tweedie, R. (2000). Trim and Fill: A Simple Funnel-Plot–Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. Biometrics, 56(2), 455-463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Elkins, A., & Zickgraf, H. F. (2018). Picky eating and food neophobia: Resemblance and agreement in parent/young adult dyads. Appetite, 126, 36-42. https://doi.org/10.1016/j.appet.2018.02.021

Fallon, A. and Rozin, P. (1983). The psychological bases of food rejection by humans. Ecology of Food and Nutrition. 13:1, 15–26, https://doi.org/10.1080/03670244.1983.9990728

Fernández-Ruiz, V., Claret, A., & Chaya, C. (2013). Testing a Spanish version of the Food Neophobia Scale. Food Quality and Preference, 28(1), 222-225. https://doi.org/10.1016/j.foodqual.2012.09.007

Frank, R. A., & van der Klaauw, N. J. (1994). The contribution of chemosensory factors to individual differences in reported food preferences. Appetite, 22(2), 101-123. https://doi.org/10.1006/appe.1994.1011

Guzeller, C.O., & Celiker, N. (2020). Examining the relationship between organizational commitment and turnover intention via a meta-analysis. International Journal of Culture, Tourism and Hospitality Research, 14(1), 102-120. https://doi.org/10.1108/IJCTHR-05-2019-0094

Hartmann, C., & Siegrist, M. (2018). Development and validation of the Food Disgust Scale. Food Quality and Preference, 63, 38-50. https://doi.org/10.1016/j.foodqual.2017.07.013

IntHout, J., Ioannidis, J. P. A., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open, 6(7), e010247. https://doi.org/10.1136/bmjopen-2015-010247

Jaeger, S. R., Rasmussen, M. A., & Prescott, J. (2017). Relationships between food neophobia and food intake and preferences: Findings from a sample of New Zealand adults. Appetite, 116, 410–422. https://doi.org/10.1016/j.appet.2017.05.030

Kaiser, L. L., Schneider, C., Mendoza, C., George, G., Neelon, M., Roche, B., & Ginsburg, D. (2012). Development and use of an evaluation tool for taste-testing activities by school-aged children. Journal of the Academy of Nutrition and Dietetics, 112(12),2028-2034. https://doi.org/10.1016/j.jand.2012.07.006

Koivisto Hursti, U. K., & Sjödén, P. O. (1997). Changing food habits in children and adolescents Experiences from intervention studies. Näringsforskning, 41(1), 102-110. https://doi.org/10.3402/fnr.v41i0.1755

Konaklıoğlu, E., & Algül, F. (2022). Turistlerin gıda neofobisi ve gıda çeşitlilik arayışının destinasyonu tekrar ziyaret etme niyeti ile ilişkisinin belirlenmesi. Ankara Hacı Bayram Veli Üniversitesi Turizm Fakültesi Dergisi, 25(1), 120-144. https://doi.org/10.55931/ahbvtfd.1128930

Lafraire, J., Rioux, C., Giboreau, A., & Picard, D. (2016). Food rejections in children: Cognitive and social/environmental factors involved in food neophobia and picky/fussy eating behaviour. Appetite, 96, 347–357. https://doi.org/10.1016/j.appet.2015.09.008

Laureati, M., Spinelli, S., Monteleone, E., Dinnella, C., Prescott, J., Cattaneo, C., ... & Pagliarini, E. (2018). Associations between food neophobia and responsiveness to "warning" chemosensory sensations in food products in a large population sample. Food Quality and Preference, 68, 113-124. https://doi.org/10.1016/j.foodqual.2018.02.007

Meiselman, H. L., King, S. C., & Gillette, M. (2010). The demographics of neophobia in a large commercial US sample. Food Quality and Preference, 21(7), 893–897. https://doi.org/10.1016/j.foodqual.2010.05.009

Monneuse, M. O., Rigal, N., Frelut, M. L., Hladik, C. M., Simmen, B., & Pasquet, P. (2008). Taste acuity of obese adolescents and changes in food neophobia and food preferences during a weight reduction session. Appetite, 50(2-3), 302-307. https://doi.org/10.1016/j.appet.2007.08.004

Morris, S. B. (2023). Meta-Analysis in Organizational Research: A Guide to Methodological Options. Annual Review of Organizational Psychology and Organizational Behavior, 10(1), 225-259. https://doi.org/10.1146/annurev-orgpsych-031921-021922

Murray, J. M., Easton, K., & Best, D. J. (2001). A Study Of Chinese-Origin And European-Origin Australian Consumers'texture Preferences Using A Novel Extruded Product. Journal of

sensory studies, 16(5), 485-504. https://doi.org/10.1111/j.1745-459X.2001.tb00315.x

Nordin, S., Broman, D. A., Garvill, J., & Nyroos, M. (2004). Gender differences in factors affecting rejection of food in healthy young Swedish adults. Appetite, 43(3), 295–301. https://doi.org/10.1016/j.appet.2004.07.002

Nunez-Nunez, R. M., Rubio-Aparicio, M., Marín-Martínez, F., Sánchez-Meca, J., López-Pina, J. A., & López-López, J. A. (2022). A Reliability Generalization Meta-Analysis of the Padua Inventory-Revised (PI-R). International Journal of Clinical and Health Psychology, 22(1), 100277. https://doi.org/10.1016/j.ijchp.2021.100277

Nunnally, J., & Bernstein, I. (1994). Psychometric theory (3rd edition). McGraw-Hill.

Özdemir, V., Yıldırım, Y., & Tan, Ş. (2020). A Meta-Analytic Reliability Generalization Study of the Oxford Happiness Scale in Turkish Sample. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 11(4), 374-404.

Öztürk, G. (2019). Etnik gıdaya yönelik tutum ile kültürleşme, kültürel açıklık ve neofobi'nin ilişkisi: Samsun Gürcü mutfağı örneği. Yüksek Lisans Tezi, Gaziantep Üniversitesi Sosyal Bilimler Enstitüsü.

Paupério, A., Severo, M., Lopes, C., Moreira, P., Cooke, L., & Oliveira, A. (2014). Could the Food Neophobia Scale be adapted to pregnant women? A confirmatory factor analysis in a Portuguese sample. Appetite, 75, 110-116. https://doi.org/10.1016/j.appet.2013.12.023

Phillips, C. E., King, C., Kivisalu, T. M., & O'Toole, S. K. (2016). A Reliability Generalization of the Suinn-Lew Asian Self-Identity Acculturation Scale. SAGE Open, 6(3), 2158244016661748. https://doi.org/10.1177/2158244016661748

Pliner, P., & Hobden, K. (1992). Development of a scale to measure the trait of food neophobia in humans. Appetite, 19(2), 105-120. https://doi.org/10.1016/0195-6663(92)90014-W

Pliner, P., Pelchat, M., & Grabski, M. (1993). Reduction of neophobia in humans by exposure to novel foods. Appetite, 20(2), 111-123. https://doi.org/10.1006/appe.1993.1013

Polat, E., & Koseoglu, M. A. (2023). Wellness tourism scholarship: A research agenda. Journal of Hospitality and Tourism Insights, 6(5), 1660-1689. https://doi.org/10.1108/JHTI-05-2022-0186

Rabadán, A., Alvarez-Ortí, M., Martínez, E., Pardo-Gim´enez, A., Zied, D. C., & Pardo, J. E. (2021). Effect of replacing traditional ingredients for oils and flours from nuts and seeds on the characteristics and consumer preferences of lamb meat burgers. LWT, 136, 110307. https://doi.org/10.1016/j.lwt.2020.110307

Rabadán, A., & Bernabéu, R. (2022). A systematic review of studies using the Food Neophobia Scale: Conclusions from thirty years of studies. Food Quality and Preference, 93, 104241. https://doi.org/10.1016/j.foodqual.2021.104241

Raudenbush, B., Van Der Klaauw, N. J., & Frank, R. A. (1995). The contribution of psychological and sensory factors to food preference patterns as measured by the Food Attitudes Survey (FAS). Appetite, 25(1), 1-15. https://doi.org/10.1006/appe.1995.0037

Reverdy, C., Chesnel, F., Schlich, P., K¨oster, E. P., & Lange, C. (2008). Effect of sensory education on willingness to taste novel food in children. Appetite, 51(1), 156–165. https://doi.org/10.1016/j.appet.2008.01.010

Ritchey, P. N., Frank, R. A., Hursti, U.-K., & Tuorila, H. (2003). Validation and cross-national comparison of the food neophobia scale (FNS) using confirmatory factor analysis. Appetite, 40(2), 163-173. https://doi.org/10.1016/S0195-6663(02)00134-4

Rosenberg, M. S. (2005). The file-drawer problem revisited: A general weighted method for calculating fail-safe numbers in meta-analysis. Evolution; International Journal of Organic Evolution, 59(2), 464-468.

Rosenthal, R. S. (1979). Release of soluble peptidoglycan from growing gonococci: Hexaminidase and amidase activities. Infection and Immunity, 24(3), 869-878. https://doi.org/10.1128/iai.24.3.869-878.1979

Rubio, B., Rigal, N., Boireau-Ducept, N., Mallet, P., & Meyer, T. (2008). We are measuring willingness to try new foods: A self-report questionnaire for French-speaking children. Appetite, 50(2-3), 408-414. https://doi.org/10.1016/j.appet.2007.09.012

Sanchez-Meca, J., Marín-Martínez, F., López-López, J. A., Núñez-Núñez, R. M., Rubio-Aparicio, M., López-García, J. J., López-Pina, J. A., Blázquez-Rincón, D. M., López-Ibáñez, C., & López-Nicolás, R. (2021). Improving the reporting quality of reliability generalisation meta-analyses: The REGEMA checklist. Research Synthesis Methods, 12(4), 516-536. https://doi.org/10.1002/jrsm.1487

Semma, B., Henri, M., Luo, W., & Thompson, C. G. (2019). Reliability Generalization of the Meaning in Life Questionnaire Subscales. Journal of Psychoeducational Assessment, 37(7), 837–851. https://doi.org/10.1177/0734282918800739

Sen, S. (2022). A Reliability Generalization Meta-Analysis of Runco Ideational Behavior Scale. Creativity Research Journal, 34(2), 178-194. https://doi.org/10.1080/10400419.2021.1960719

Shou, Y., & Olney, J. (2020). Assessing a domain-specific risk-taking construct: A meta-analysis of reliability of the DOSPERT scale. Judgment and Decision Making, 15(1), 112–134. https://doi.org/10.1017/S193029750000694X

Siegrist, M., Hartmann, C., & Keller, C. (2013). Antecedents of food neophobia and its association with eating behaviour and food choices. Food Quality and Preference, 30(2), 293-298. https://doi.org/10.1016/j.foodqual.2013.06.013

Sivrikaya, K. K., & Pekerşen, Y. (2020). The impact of food neophobia and sensation seeking of foreign tourists on the purchase intention of traditional Turkish food. International Journal of Gastronomy and Food Science, 21, 100222. https://doi.org/10.1016/j.ijgfs.2020.100222

Skinner, J. D., Carruth, B. R., Bounds, W., & Ziegler, P. J. (2002). Children's food preferences: a longitudinal analysis. Journal of the American Dietetic Association, 102(11), 1638-1647. https://doi.org/10.1016/S0002-8223(02)90349-4

Sogari, G., Menozzi, D., & Mora, C. (2019). The food neophobia scale and young adults' intention to eat insect products. International Journal of Consumer Studies, 43(1), 68-76. https://doi.org/10.1111/ijcs.12485

Tuorila, H., & Mustonen, S. (2010). Reluctant trying of an unfamiliar food induces negative affection for the food. Appetite, 54(2), 418-421. https://doi.org/10.1016/j.appet.2010.01.010

Uçar, E. M., Gümüş, D., Karabulut, E., & Kızıl, M. (2021). Yeni besin korkusu ölçeğinin Türkçeye uyarlaması ve uygun faktör yapısının belirlenmesi. Turkiye Klinikleri Journal of Health Sciences, 6(3). https://doi.org/10.5336/healthsci.2020-76664

Üngüren, E., & Tekin, Ö. A. (2022). The effect of openness to experience personality trait of kitchen staff on creativity potential: The mediating effect of food neophobia and the moderating effect of occupational self-efficacy. International Journal of Gastronomy and Food Science, 28, 100530. https://doi.org/10.1016/j.ijgfs.2022.100530

Vacha-Haase, T. (1998). Reliability Generalization: Exploring Variance in Measurement Error Affecting Score Reliability Across Studies. Educational and Psychological Measurement, 58(1), 6-20. https://doi.org/10.1177/0013164498058001002

Vassar, M. (2008). A note on the score reliability for the Satisfaction with Life Scale: An RG study. Social Indicators Research, 86(1), 47–57. https://doi.org/10.1007/s11205-007-9113-7

Yörük, S., & Sen, S. (2022). A Reliability Generalization Meta-Analysis of the Creative Achievement Questionnaire. Creativity Research Journal, 0(0), 1-16. https://doi.org/10.1080/10400419.2022.2148073

Zhao, J., Gao, Z., Li, Y., Wang, Y., Zhang, X., & Zou, L. (2020). The food neophobia scale (FNS): Exploration and confirmation of factor structure in a healthy Chinese sample. Food Quality and Preference, 79, 103791. https://doi.org/10.1016/j.foodqual.2019.

SUBMITTED: JUL 2023
REVISION SUBMITTED: OCT 2023
ACCEPTED: JAN 2024
REFEREED ANONYMOUSLY

PUBLISHED ONLINE: 30 JUNE 2024