CrossMark

Undisguised disgust: a psychometric evaluation of a disgust propensity measure

Jakub Polák[1,2] · Eva Landová[1] · Daniel Frynta[1]

Abstract

Disgust, considered as one of the basic emotions universally shared across various cultures, has only recently attracted increased research interest. Originally developed as a key component of the behavioural immune system serving to protect the body from infectious pathogens, disgust has taken on other roles such as guiding our moral judgements, political preferences, or even choices of sexual partners. Despite these adaptive roles, disgust may, on the other hand, be responsible for negative social (xenophobia, homophobia) or clinical phenomena (common anxiety disorders such as specific phobias of obsessive-compulsive disorder), thus, a reliable measure of individual tendencies to experience disgust is indispensable. The aim of this study was to develop and analyse a Czech version of the Disgust Scale – Revised, one of the most widely used assessment of disgust propensity. Using a back-translation procedure and a counterbalanced test-retest experimental design, we demonstrated that scores on the Czech measure correlated with those on the English original (test-retest reliability $r = 0.82$) and both measures were found equivalent ($p = 0.019$). We also adopted a general linear model to show that the level of disgust is affected by sex, age, and education/occupancy, as women and people with no biology education score significantly higher than men and biologists, respectively. Finally, our data provide support to a bifactor model composed of general disgust taping into all items in addition to three distinct disgust domains.

Keywords Behavioral immune system · Contamination · Disgust scale · DS-R · Test translation

Introduction

Within the growing body of research on human emotions, disgust has only recently attracted substantial interest of various disciplines (Rozin et al. 2016). Nowadays, it is considered as one of the basic emotions with a cross-culturally recognizable facial expression (Ekman 1992), distinct physiological response (Stark et al. 2005) and specific neuroanatomical correlates (reviewed in Schienle 2009). Originally, it serves as the "behavioural immune system" protecting the body from dangerous pathogens (Curtis et al. 2011). However, through a

process of preadaptation disgust has been useful in serving new functions, such as guiding moral judgments (Schnall et al. 2008) or facilitating the in-group cohesion through xenophobia (Faulkner et al. 2004). Disgust is also of interest to clinical researchers due to its role in aetiology of various mental disorders ranging from animal phobias (Davey and Marzillier 2009), contamination-based obsessive-compulsive disorder (OCD), eating disorders, and sexual dysfunctions (Olatunji et al. 2010).

Despite the need of a reliable valid measurement of disgust (Olatunji and Sawchuk 2005), it has become a challenging task given the vast diversity of disgust elicitors that extends beyond contaminated food (Rozin et al. 2016). Waste products, body envelope violations (i.e. surgery, puncture wounds, deformity, and other situations in which the normal exterior envelope of the body is breached or altered, Haidt et al. 1994), deviant sexual practices, or even violations of adopted moral codes (Chapman et al. 2009; Tybur et al. 2013) are new sources of disgust that require an inclusive and reliable form of measurement. Haidt et al. (1994) put forth an influential theoretical model categorizing disgust into two principal subtypes: core disgust (food, animals, and body products) and

✉ Jakub Polák
   polak.jakub@seznam.cz

1 National Institute of Mental Health, Topolová 748, 250 67 Klecany, Czech Republic

2 Department of Psychology, Faculty of Arts, Charles University, Celetná 20, 116 42 Prague 1, Czech Republic

animal reminder disgust (death, sex, body envelope violations). To measure disgust propensity (individual tendency to experience disgust) the authors presented a self-report instrument, the Disgust Scale (DS), which has become the method of choice with more than a hundred published articles (Olatunji and Sawchuk 2005). Recently, Olatunji et al. (2007) suggested to revise the measure by dropping seven items and regroup the rest into three factors called core, animal reminder, and contamination-based disgust.

Although the Disgust Scale – Revised (DS-R) appears to have some utility across cultures (Olatunji et al. 2009), a specific cultural background can modulate what objects or situations are considered being disgust elicitors (Rozin and Fallon 1987; Tybur et al. 2009). It is thus valuable to possess a local instrument for measuring disgust propensity that allows cross-national comparisons. So far, several language versions of the DS-R have been developed (e.g., Dutch, Korean, or Hebrew), but a Czech translation has yet to come.

Therefore, the aim of this study was to (1) create a standardized Czech translation of the DS-R, (2) analyse its psychometric qualities including reliability and validity (using an alternative measure of disgust propensity, the Three-Domain Disgust Scale, TDDS; Tybur et al. 2009), (3) verify its three-factor structure, and (4) test the effect of selected demographic variables (sex, age, education) on disgust scores. As people from different disciplines are exposed to varying levels of disgust elicitors, which require a different ability to manage these experiences, we expect people with biology education to demonstrate lower disgust propensity than those educated in social sciences (psychology, sociology, law, etc.).

Methods

Subjects

In total, 1069 healthy subjects of both sexes (272 men and 797 women) aged 16–89 years were recruited for the study. Most of them were high school ($N = 81$) or college students ($N = 751$). In order to test for the effect of study field, students of natural (biology) and social sciences (psychology, sociology, politics, etc.) took part. The additional participants were recruited from staff members of a biology faculty and psychiatric research centre, or through the Internet. The study has been approved by the respective institutional ethics committee. All the subjects gave their written informed consent prior to testing.

Psychometrics

The DS-R is a self-report personality scale to assess individual differences in propensity to disgust developed by Haidt et al.

(1994) and modified by Olatunji et al. (2007). There are 25 disgust elicitor items loading on one of the three factors (core disgust, animal reminder disgust, contamination-based disgust) and two catch questions (item 12 and 16) to identify those respondents that are not paying attention to the task or do not take it seriously. Each item is rated by the participant on a 5-point Likert scale from 0 ("Strongly disagree/Not disgusting at all") to 4 ("Strongly agree/Extremely disgusting"). The total score (ranging from 0 to 100) is calculated by summing scores on all the 25 disgust elicitor items but three (item 1, 6, 10) that are reverse-scored. Similarly, subscale scores may be calculated. All the participants that do not give valid answers on the catch questions should be dropped. The DS-R demonstrates acceptable Cronbach's alpha estimates for the overall internal consistency (0.84) and the three subscales (core disgust: 0.74; animal reminder disgust: 0.78; contamination-based disgust: 0.61; Olatunji et al. 2007, which was replicated by van Overveld et al. 2011).

The TDDS is a 21-item self-report measure of disgust propensity in three domains: pathogen, sexual, and moral disgust. The items are scored on a 7-point Likert-type scale. The three subscales have demonstrated very good internal consistency (Cronbach's alpha: 0.84, 0.87, and 0.84, for pathogen, sexual, and moral disgust, respectively) and validity (Olatunji et al. 2012; Tybur et al. 2009, 2011). Prior to the study it has been translated into Czech by two independent bilingual translators.

Procedures

The standardization procedure followed the guidelines for translating and adapting tests set by the International Test Commission (ITC 2017; see also Hambleton et al. 2005). First, having the permission of the copyright holder, the original DS-R was translated from English to Czech independently by two bilingual professionals. These two versions were then checked by a psychologist experienced in test development to identify and resolve potential item discrepancies in the translations. Subsequently, a back-translation to English was performed by another translator unfamiliar with the questionnaire. Three native English speakers then compared the original and back-translated items to determine whether they were equivalent in meaning. Any substantive differences in particular items were considered and appropriately revised by a team of psychologists with the objective to obtain a translation best corresponding the original instrument.

Next, a counterbalanced experimental design was adopted to standardize the Czech translation of the DS-R and evaluate its psychometric properties. Thus, a half of the subjects were administered the original English version first, followed by the Czech one approximately two months later. The other half was asked to complete the questionnaires in the reverse order,

i.e., first in Czech and then in English. The way the participants were divided into these two groups was completely random. The selected period of two months between each administration is generally recommended when retesting personality questionnaires (McCrae et al. 2011). It is believed that after this time the subjects can no longer remember their previous answers that could influence the current score, thus the carry over effect is eliminated (McConnell et al. 1998). In addition, to check for convergent validity, a subsample of 288 respondents also completed the Czech version of the TDDS a few months later.

Prior to testing, we asked the subjects about their English proficiency and instructed them not to complete the instrument if they did not feel confident. However, as mentioned above, the majority of our sample were high school or university students who have learned English for several years. Moreover, as a part of the curriculum at Charles University, students are required to pass an English exam equivalent to "A Cambridge English: First" qualification (FCE). Therefore, a good level of English skills among our participants was expectable. Information on the individual's age, sex, and occupancy/education was collected as well. For statistical analyses, the age was categorized into seven groups (less than 15; 15–20; 21–30; 31–40; 41–50; 51–60; and more than 60 years). The occupancy/education was categorized either as 1) biological (university students of biology, lecturers of the Faculty of Science, and doctors) and 2) other than biological (high school students, university students of psychology, sociology, economics, politics, and law, IT technicians, psychologists, etc.).

Statistical Analyses

First, the scores distribution from both language versions was checked for normality using a Shapiro-Wilk test. Based on the outcome, we employed either a paired samples t-test or Wilcoxon signed-ranked test to analyse the effect of administration order (test vs. retest) and language (Czech vs. English). Responses on individual items were compared using a Wilcoxon matched-pair signed-rank test for related samples. Subsequently, a score difference between the test and retest was calculated and a t-test for independent samples was employed to check, whether this was affected by the language of the questionnaire used first. Score correlation between both language versions was calculated as well using a formula for test-retest reliability.

We also employed methods of statistical equivalence testing, specifically a two one-sided t-test (TOST; Schuirmann 1987), to analyse the measurement invariance. An acceptance criterion θ was calculated using the following formula: $\theta = \delta + s' \left[t_{(1-\alpha,2n-2)} + t_{(1-\beta/2,2n-2)} \right] \sqrt{\frac{2}{n}}$ (δ: the absolute value of true

difference between the groups' mean values, arbitrarily set to 0; s': the upper 95% confidence interval (CI) of standard deviation s; t-test critical values for $\alpha = \beta = 0.05$). Subsequently, a 90% CI of difference in mean total scores from both measures was compared to $[-\theta; \theta]$ interval.

Secondly, to verify psychometric properties of the translated instrument, total scores from the Czech DS-R were analysed separately. The items factor structure was examined using a factor analysis with the maximum likelihood extraction method and Direct Oblimin rotation that allows for correlation which is expectable in psychological phenomena (Costello and Osborne 2005). As the data were on a 5-point Likert scale, they were treated as ordinal. The appropriate number of factors to retain was determined using a parallel analysis (Horn 1965) which has been demonstrated as one of the most accurate methods for such purpose (Zwick and Velicer 1986). Here we followed a procedure developed by O'Connor (2000) using his updated syntax (O'Connor n.d.). We ran 5000 random permutations of a raw data set with 891 cases and 25 variables (i.e. the disgust elicitor items; item 1, 6, and 10 were first rescored) to which the real data eigenvalues were compared. Furthermore, in order to verify the factor structure of the Czech DS-R, Tucker's congruence coefficients of factor similarity with the original scale were calculated and the same analysis was repeated after a Procrustes rotation had been applied to our matrix of factor loadings. We have also computed a Mantel test to compare the items correlation matrix calculated from our data with that originally published by Olatunji et al. (2007).

The model fit was then checked by conducting a confirmatory factor analysis (CFA) comparing several alternative factor solutions based on the literature (including a parsimonious one-factor model and a bifactor model allowing to test for a general latent "g" disgust factor in addition to distinct components of disgust; see Olatunji et al. 2014). Competing nested models were compared using a chi-square difference test (χ^2_{diff}).

Reliability was calculated using the split-half method and items internal consistency was expressed as the Cronbach's alpha. We also used the Pearson correlation coefficient between the three subscale scores of the DS-R and TDDS to demonstrate convergent validity. Finally, a General Linear Model (GLM) was used to analyse the effect of age, sex, and education background. Most of the calculations were performed in the SPSS Statistics, version 22 (IBM Corp. 2013). The CFA was conducted in the SPSS Amos, version 24 (Arbuckle 2016). Finally, the Mantel test and Tucker's congruence coefficients were calculated in the XLSTAT add-on statistical package for Excel, version 2017.4 (Addinsoft 2017).

Results

Translation Evaluation

In total, 351 participants completed the DS-R in both languages, the rest of subjects did not take part in retest (62 completed the DS-R in English, 656 in Czech). The fit of total scores on both versions to a normal distribution was acceptable at $\alpha = 0.01$ (English: $p = 0.014$, skewness 0.15, kurtosis −0.58; Czech: $p = 0.032$, skewness 0.18, kurtosis −0.53). We found a statistically significant difference in total scores on the English ($M = 44.38 \pm 0.77$, $SD = 14.46$) and Czech DS-R ($M = 42.38 \pm 0.72$, $SD = 13.51$); $t(350) = 4.44$, $p < 0.001$; see Table 1). Item analyses with a Bonferroni correction revealed that responses on both versions differed significantly at the selected level ($p < 0.002$) on 7 items (1, 2, 6, 10, 18, 24, and 27).

We also found a slight though nonsignificant decrease in total scores from the first ($M = 43.73 \pm 0.75$, $SD = 14.01$) to the second administration ($M = 43.03$, $SD = 14.03$); $t(350) = 1.52$, $p = 0.128$. The test-retest score difference was significantly higher when the English DS-R was administered first ($M = 3.15 \pm 0.76$, $SD = 9.25$) compared to the reversed order ($M = -1.13 \pm 0.54$, $SD = 7.63$); $t(349) = 4.75$, $p < 0.001$. Based on these results, the studied sample was divided respective to the version tested first and each data subset was analysed separately.

When the original instrument was tested first, it yielded a significantly higher total score ($M = 47.94 \pm 1.25$, $SD = 15.24$) compared to retest using the Czech translation ($M = 44.79 \pm 1.22$, $SD = 14.88$); $t(149) = 4.18$, $p < 0.001$. Interestingly, the opposite pattern, although less pronounced, was found when the Czech DS-R was tested prior to the original ($M = 40.59 \pm 0.85$, $SD = 12.11$ vs $M = 41.79 \pm 0.94$, $SD = 13.25$); $t(200) = -2.10$, $p = 0.037$ (see Fig. 1). Despite these effects, test-retest reliability remained relatively high for the total score ($r = 0.82$, $p < 0.01$; see Table 1 for more results).

Additionally, two alternative methods of the TOST confirmed that the two instruments were equivalent in measuring disgust propensity. The 90% CI of difference in mean total scores ($M = 1.99 \pm 1.74$) was completely contained within the interval of acceptance criterion (−4.20; 4.20). This was supported by the right and left one-sided t-test on the lower and upper bound; $t(700) = -2.09$, $p = 0.019$ and $t(700) = 5.87$, $p < 0.001$, respectively.

The following analyses were performed on the data from the Czech DS-R only ($N = 1006$). There was no significant difference in total scores between those who completed the translated instrument first ($N = 857$, $M = 43.27 \pm 0.49$, $SD = 14.34$) or second ($N = 149$, $M = 44.71 \pm 1.22$, $SD = 14.90$); $t(1004) = -1.13$, $p = 0.260$), leading to all data being pooled together. Considerably lower scores were found in subjects who provided invalid responses on the catch items 12 and 16 ($N = 115$, $M = 39.88 \pm 1.38$, $SD = 14.80$ vs. $N = 891$, $M = 43.95 \pm 0.48$, $SD = 14.31$); $t(1004) = 2.85$, $p = 0.004$, thus, they were excluded from further analyses.

Factor Analysis

The Kaiser-Meyer-Olkin measure of sampling adequacy (0.88) and Bartlett's test of sphericity ($\chi^2 = 4258.70$, $df = 300$, $p < 0.001$) confirmed that the item structure of the data set warranted a factor analysis. Although six factors had eigenvalues greater than one, the parallel analysis revealed that only the first four eigenvalues extracted from the real data exceeded both the mean and the 95th percentile of those based on the simulations of raw data sets (5.50, 1.76, 1.36, and 1.24). However, a close visual inspection of the scree plot showed that the fourth eigenvalue was not significantly different from the 95th percentile of the corresponding random data eigenvalue (1.24 vs. 1.23) and just represented a correlated residual of the items (Panter et al. 1997). Thus, the

Table 1 Comparison of results from the English and Czech version of the Disgust Scale – Revised (DS-R)

	Lang.	Mean	SD	SE	r	t	df	p
Total score	EN	44.38	14.46	0.77	0.82	4.44	350	< 0.001
	CZ	42.38	13.51	0.72				
Core	EN	23.56	7.30	0.39	0.80	4.49	350	< 0.001
	CZ	22.45	7.14	0.38				
Animal reminder	EN	14.02	6.12	0.33	0.76	0.22	350	0.826
	CZ	13.97	6.39	0.34				
Contamination	EN	6.79	3.96	0.21	0.71	5.52	350	< 0.001
	CZ	5.97	3.21	0.17				

Only those respondents who completed both instruments were included in the analysis ($N = 351$)

Lang. language version of the DS-R, *EN* English, *CZ* Czech, *SD* standard deviation, *SE* standard error of mean, *r* correlation coefficient between total/subscale scores from both language versions, *t* t-test value comparing total/subscale scores from both language versions, *df* degrees of freedom
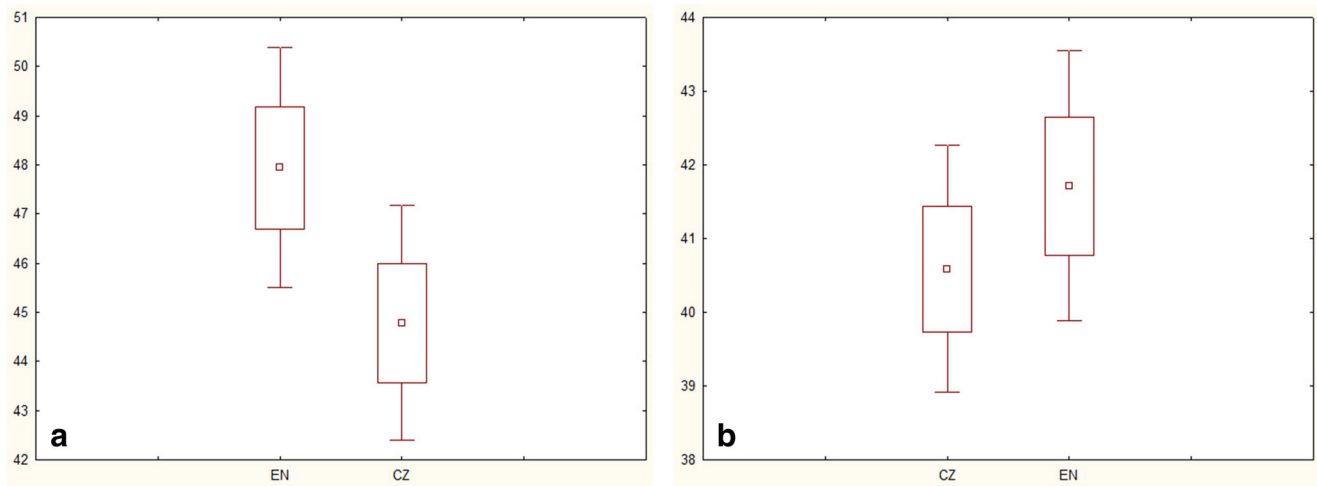
Fig. 1 Comparison of the total scores from the English and Czech DS-R. **a** The English original administered before the Czech translation; **b** The Czech translation administered before the English original. Inner square = mean; box = mean ± SE; whiskers = mean ± 1.96*SE

parallel analysis revealed that only three factors should be extracted, two of which were clear, while the third one was weaker, though still salient.

This three-factor solution explained 25.96% of the total variance. The first factor accounted for 19.04% of the total variance and consisted of nine items, all of them tapping into animal reminder with only one item from the core disgust subscale, though its factor loading was very low (0.202). The second factor, which accounted for 4.45% of the total variance, included nine items scoring on the core and one on the contamination-based disgust subscale. The last factor accounting for 2.47% of the total variance then contained six items, four belonging to the contamination-based and two to the core disgust subscale (see Table 2 for factor loadings).

Despite a relatively weak structure of the third factor, Tucker's coefficients of congruence with the original factor matrix were satisfactory for all three factors ($r = 0.90$, 0.86, and -0.81) and these slightly improved after the Procrustes rotation ($r = 0.91$, 0.86, and 0.82). This was corroborated by the Mantel test that revealed a significant correlation between the items correlation matrices ($r = 0.73$; $p < 0.001$).

Although there was a reasonable fit of the three-factor model ($\chi^2 (272) = 889.40$, $p < 0.001$) as revealed by the CFA using various indices, including the Root Mean Square Error of Approximation (RMSEA = 0.05), Comparative Fit Index (CFI = 0.846), and the Tucker-Lewis Index (TLI = 0.830), the best fit among all the models was associated with the bifactor model of general disgust and three distinct disgust domains ($\chi^2 (250) = 739.04$, $p < 0.001$; RMSEA = 0.047; CFI = 0.878; TLI = 0.854). Thus, adding a "g" disgust factor with loadings from all the items significantly improved the model ($\chi^2_{diff} (22) = 150.35$, $p < 0.001$). Conversely, neither

Table 2 Factor loadings of the disgust elicitor items from the Czech DS-R. Extraction method: maximum likelihood, rotation method: Oblimin with Kaiser normalization ($N = 892$)

Item	Subscale	Factor		
		1	2	3
7	Animal reminder	**0.813**	−0.116	−0.034
2	Animal reminder	**0.628**	−0.015	−0.018
21	Animal reminder	**0.563**	0.231	0.169
19	Animal reminder	**0.477**	0.117	−0.130
24	Animal reminder	**0.453**	0.133	−0.239
14	Animal reminder	**0.420**	0.020	−0.222
10[a]	Animal reminder	**0.373**	0.137	0.054
5	Animal reminder	**0.347**	0.000	−0.254
11	Core	**0.202**	0.116	−0.173
25	Core	−0.159	**0.535**	−0.118
15	Core	0.183	**0.482**	0.118
17	Core	−0.062	**0.443**	−0.205
27	Core	0.082	**0.431**	−0.112
3	Core	0.102	**0.408**	0.076
20	Core	−0.013	**0.389**	−0.075
8	Core	0.149	**0.367**	0.128
22	Core	−0.040	**0.338**	−0.284
23	Contamination	0.034	**0.264**	−0.209
6[a]	Core	0.081	**0.254**	−0.019
13	Core	0.115	−0.005	**−0.458**
4	Contamination	0.011	−0.018	**−0.449**
9	Contamination	−0.026	0.062	**−0.398**
1[a]	Core	0.155	0.011	**−0.355**
18	Contamination	−0.003	0.147	**−0.340**
26	Contamination	0.093	0.185	**−0.247**

[a] Items have been re-scored prior to the analysis (see the Methods). The highest factor loading for each item is presented in boldface

Table 4 Spearman correlation coefficients between the total and subscale scores on the Czech version of the DS-R (reflecting the item structure as shown in Table 2) and TDDS ($N = 288$)

Disgust domain	1	2	3	4	5	6	7	8
1. DS-R total score	–	0.850[**]	0.823[**]	0.693[**]	0.345[**]	0.526[**]	0.293[**]	0.020
2. Factor I		–	0.516[**]	0.411[**]	0.239[**]	0.381[**]	0.209[**]	0.006
3. Factor II			–	0.453[**]	0.311[**]	0.547[**]	0.259[**]	−0.030
4. Factor III				–	0.345[**]	0.364[**]	0.305[**]	0.102
5. TDDS total score					–	0.490[**]	0.733[**]	0.732[**]
6. Pathogen						–	0.307[**]	−0.009
7. Sexual							–	0.276[**]
8. Moral								–

[**] Correlation is significant at $p < 0.001$ (two-tailed)

the unidimensional model with 25 indicator variables, nor a two-factor model showed a good fit to the data (see Table 3 for more results from the CFA).

Psychometric Properties of Czech DS-R

The total scores were normally distributed (skewness 0.15; kurtosis −0.42); see a table in Online Resource 1 for norms. The disgust elicitor items showed a moderate reliability through examining the split-half Spearman-Brown coefficient ($r = 0.70$) or the Cronbach's alpha coefficient of internal consistency ($\alpha = 0.75$). The DS-R total scores significantly correlated with those obtained on the TDDS ($r = 0.35$, $p < 0.001$); see Table 3 for more results.

The GLM model revealed a significant effect of sex ($F(1,773) = 68.74$, $p < 0.001$), age ($F(6,668) = 8.55$, $p < 0.001$), and occupancy/education ($F(1,773) = 45.53$, $p < 0.001$) on the DS-R total score. Conversely, neither the sex*age nor sex*occupancy/education interaction effect proved significant ($F(5,666) = 1.02$, $p = 0.402$ and $F(1,771) = 3.33$, $p = 0.068$, respectively); for parameter estimates see a table in Online Resource 2. On average, women scored significantly higher than men ($M = 46.06 \pm 0.54$ vs. 36.93 ± 0.89) and people with other than biology education scored significantly higher than biology students and postgraduates ($M = 46.34 \pm 0.67$ vs. 39.42 ± 0.72; see Table 4 and 5 for more details).

Discussion

Our data, gathered on respondents who completed both the English and translated instrument, show a high correlation of the total scores ($r = 0.82$). This result is within the range of 0.7–0.9 which is recommended for test-retest reliability of psychological assessment (Groth-Marnat 2009) and corresponds well to test-retest reliability as assessed with a Spearman correlation coefficient 0.85 reported by Kang et al. (2012) who adapted the DS-R into Korean. Slightly lower though still satisfactory correlations were found in our study for the scores on three subscales, i.e. core, animal reminder, and contamination-based disgust ($r = 0.79$, 0.76, and 0.71, respectively). These values clearly demonstrate that despite a relatively long delay between both administrations, which exceeds the usually advised period of one month, responses on the DS-R are relatively consistent and the adapted Czech measure provides reliable data that correlate significantly with the original DS-R. Moreover, we have demonstrated that the two instruments may be considered equivalent ($p = 0.019$) in measuring disgust propensity.

Further statistical analyses revealed elevated total scores on the English DS-R compared to its translation irrespective of the administration order. That is if the Czech respondents tend to self-report higher disgust propensity when answering in English than in their native language. This difference in scores remains even if they complete the original measure after having

Table 3 Comparison of different models of the Czech DS-R using a confirmatory factor analysis showing selected indices of model fit ($N = 892$). The models are in the order of increasing fit

Model tested	χ^2	df	χ^2/df	AIC	AGFI	CFI	TLI	RMSEA
1-factor model	1260.60[**]	275	4.58	1360.60	0.859	0.754	0.732	0.063
2-factor model	984.77[**]	274	3.59	1086.77	0.900	0.823	0.806	0.054
3-factor model	889.40[**]	272	3.27	995.40	0.906	0.846	0.830	0.050
Bifactor model	739.04**	250	2.96	889.04	0.917	0.878	0.854	0.047

df degrees of freedom, χ^2/df a ratio of chi-square divided by the degrees of freedom, AIC Akaike information criterion, AGFI adjusted goodness-of-fit test, CFI comparative fit index, TLI Tucker-Lewis index, RMSEA root mean square error of approximation

[**] χ^2 test is significant at $p < 0.001$

Table 5 Descriptive statistics of the total scores obtained by the Czech DS-R categorized according to sex, age, and type of education/occupancy

Parameter	N	Mean	SE	95% CI
Overall	891	43.95	0.48	43.01; 44.89
Men	208	36.93	0.89	35.19; 38.68
Women	679	46.06	0.54	45.00; 47.12
< 15 yrs	5	63.60	6.27	46.18; 81.02
15–19 yrs	168	49.53	1.10	47.35; 51.71
20–29 yrs	528	43.74	0.60	42.56; 44.91
30–39 yrs	110	41.05	1.33	38.42; 43.67
40–49 yrs	41	36.71	2.21	32.25; 41.16
50–59 yrs	20	38.95	3.30	32.04; 45.86
60+ yrs	18	32.06	2.26	27.28; 36.83
Biologists	336	39.42	0.72	38.00; 40.84
Non-biologists	444	46.34	0.67	45.03; 47.66

N: Out of 1006 respondents who completed the Czech DS-R (i.e. those 351 with both languages +655 others who completed only the Czech scale but not the English original), 115 questionnaires were invalidated by responses on two catch questions 12 and 16, leaving 891 subjects for psychometric analysis of the Czech DS-R including its factor structure. Furthermore, 4 participants did not provide their gender and 111 did not state their field of study, thus not allowing us to assign them either to a "biologist" or "non-biologist" group. Therefore, the GLM model analysing the effect of sex, age, and education/occupation was based on 780 subjects. *SE* standard error of mean, *CI* confidence interval

read the Czech translation, so the effect of impaired comprehension is ruled out (see Fig. 1). It may be suggested, that it is not the order of presentation (that was nonsignificant), but rather the language it is written in that may affect the DS-R scores.

Interestingly, a very similar trend has been recently found in a study on psychometrics of the Snake Questionnaire (SNAQ), a self-report measure of snake fear (Polák et al. 2016). As the data from both questionnaires follow a different distribution, i.e. normally distributed scores of disgust, but a negative binomial function in the case of snake fear, it seems that either over-scoring in English or rather under-scoring in Czech might be a general psychometric phenomenon. Perhaps, English is seen by Czech students as a more formal, academic language used in scientific discourse, which might subsequently unconsciously lower the tendency for self-styled responses. It is also possible that the translation process has slightly shifted the items meaning despite a thorough back-translation procedure used in translation development, so disgust reported on the Czech DS-R tends to be a bit lower. Overall, the possible effect of a foreign language on self-disclosure tendencies is an interesting issue that would warrant further research.

Although the DS has been considerably improved by Olatunji et al. (2007) in terms of subscales' internal consistency and its factor structure, there is still an ongoing discussion about the number of factors of the DS-R. The original eight factors were reduced to three different but interrelated disgust

factors, i.e. core, animal reminder, and contamination-based disgust. In an extensive cross-cultural study performed by Olatunji et al. (2009), the three-factor solution best accounted for data from seven countries including Australia, Brazil, Germany, Italy, Japan, Sweden, and the USA. Only the Netherlands showed a poor fit to the model, but this was later challenged by van Overveld et al. (2011) who again confirmed three factors. These were also found in an adolescent sample of US high school students (here the factors were called contagion, mortality, and contact disgust: Kim et al. 2013) or in Israeli Jewish citizens (Berger and Anaki 2014). Moreover, the concept of three different kinds of disgust was demonstrated to have a convincing convergent and divergent validity in a study by Olatunji et al. (2008) who showed that each disgust type is distinctively associated with personality, behavioural, physiological, and clinical correlates. Conversely, data from the Korean version of the DS-R best fitted to a five-factor model (Kang et al. 2012).

Based on our data and in accordance with Olatunji et al. (2014), the best fitting is a bifactor model of disgust composed of four latent variables of which one is a general disgust factor being loaded on by all the 25 DS-R items and simultaneously there are three distinct domains of disgust that were revealed in the parallel analysis as well. These three separate disgust domains then describe in more details the variance in disgust propensity towards various categories of stimuli among subjects. It should be noted however that despite having the best fit to the data, some indices are still slightly below the recommended threshold (e.g., CFI and TLI should be >0.9) and lower than those reported by Olatunji et al. (2014) for the bifactor model. It is difficult to get a good fit with a large sample size. Perhaps the measure might also need some revision, especially items 6, 11, 23, and 27 seem causing problems to the model fit due to their low standardized regression weights. Nevertheless, an independent sample would be required to re-analyze the factor structure of such revised scale.

The first and most robust factor almost exclusively consists of items from the animal reminder subscale (see Table 2) and there is a very strong correlation between the first factor scores and the animal reminder subscale scores ($r = 0.97$, $p < 0.01$). On the other hand, the remaining two factors were considerably weaker, accounting together for only around 7% of total variance, and were mutually correlated ($r = -0.44$). This was also supported by Tucker's congruence coefficients showing a fair similarity for the first two factors, though a bit lower for the third one. Note however, that the third factor referred to as contamination-based disgust was the most problematic one already in the original psychometric evaluation (Olatunji et al. 2007). In our study, no clear distinction between the original core and contamination-based disgust subscale could be inferred since items pertaining to these dimensions were mixed in the two factors.

This may reflect the conceptual ambiguity in the original model of disgust put forth by Rozin et al. (2000) and further developed by Olatunji et al. (2007). Core disgust is, according to the authors, based on a sense of offensiveness at the prospect of oral incorporation of disgusting stimuli (e.g., tainted food, body waste products, etc.). Contamination-based disgust is, on the other hand, defined as disgust reactions based on the perceived threat of transmission of contagion (e.g., interpersonal disgust associated with hygiene). In our view and in accordance with recent evolutionary models (Curtis 2011; Davey 2011), the main function of disgust is to act as the behavioural immune system which helps protect the body from the risk of disease. Stimuli triggering core (e.g., rotting foods, waste products, and small animals) or contamination touch-based disgust (e.g., a public toilet, a sick cook in a restaurant, or a new condom) may be seen together as sources of harmful pathogens and tap into a single construct of disease avoidance (Tybur et al. 2009). As Olatunji et al. (2014) demonstrated, core and contamination disgust are more strongly associated, whereas animal-reminder disgust may be more distinct. Thus, the motive to separate the core and contamination-based disgust items into two distinctive domains remains questionable (cf. van Overveld et al. 2011).

Some support for those conclusions can be found in our analysis of convergent validity. Correlation coefficients between scores on the DS-R and TDDS are significant and correspond to previous reports (Olatunji et al. 2012; Tybur et al. 2009). The strongest correlation was found between similar constructs, i.e. pathogen disgust on one hand and the three subscales of the DS-R on the other. This corroborates the view of Tybur et al. (2009) that the three components of disgust as measured by the DS-R, despite forming separate factors, conceptually all pertain to fear of pathogens transmission. Conversely, the lowest and non-significant correlations were found between moral disgust (TDDS) and the DS-R subscales. This is in line with the previous evidence (Tybur et al. 2009) and can be considered as a sign of discriminant validity as moral disgust is not being covered at all by the DS-R.

Despite the relatively abundant literature reporting psychometric properties of the DS-R, only a few studies use the new scoring format suggested by van Overveld et al. (2011; i.e. a unified 5-point Likert scale on all items vs. a true/false response on 13 items and a 3-point Likert scale on 12 items used previously). The mean total score we found on the Czech DS-R ($M = 43.95$) is considerably lower compared to the Dutch ($M = 50.45$; van Overveld et al. 2011) or Korean sample ($M = 54.92$; Kang et al. 2012). It is hard to speculate what might account for this difference. Even if we accept that diverse cultural background would be an intuitive explanation for the comparison of disgust propensity in the Czech Republic and South Korea, this is not the case of the Netherlands that is culturally very similar to Czechia.

Nevertheless, the Dutch sample was very uniform as it consisted exclusively of university students of psychology. On the other hand, we recruited people with a varied educational background (social vs. natural sciences), which was found to have a significant effect on disgust propensity. To our knowledge, this is the first time it has been tested and it provides support for construct validity of the translated questionnaire. As we show here, biologists demonstrate significantly lower disgust propensity compared to psychologists, sociologists, economists, and others. This is no surprise considering that disgust elicitors (faeces, bodily fluids, dead bodies and their parts, small animals, etc.) are at the very essence of biological science. However, similarly to what was already discussed for lower fear of snakes in biology students (Polák et al. 2016; but see Landová et al. 2018), it remains unclear whether higher disgust propensity precludes people from studying biology or rather pursuing a biology degree may lower the initial disgust propensity level through a process based on desensitization. In any case, even though the considerable number of biologists among our participants lowered the mean score, it cannot explain the difference in comparison to the Dutch study as the DS-R score in the non-biologists group was still approximately 5 points lower.

Lastly, we verified the effect of sex on disgust scores. It is generally agreed that disgust sensitivity and propensity is gender-dependent. On average, women score significantly higher than men, thus, tend to react with disgust more frequently and value this experience more negatively (Berger and Anaki 2014; Haidt et al. 1994; Rozin et al. 2016; Tybur et al. 2011). Based on our data, the DS-R score of women is more than 25% higher than that of men. Surprisingly, as has been shown in an fMRI study, higher experienced disgust in women is not reflected by increased brain activity in neuroanatomical structures processing repulsive stimuli (e.g. amygdala, insula, fusiform gyrus, or orbitofrontal cortex; Schienle et al. 2005). The effect of sex on self-reported disgust also fits into the well-known fact that women show generally higher tendency to experience negative emotions such as fear (see a review by McLean and Anderson 2009) which may develop more into some anxiety disorders. For example, up to 91% of animal phobics are women (LeBeau et al. 2010; for the link between disgust and animal phobias see a review by Davey and Marzillier 2009). From the evolutionary perspective, these results are not surprising given the higher reproductive value of women than men because of their ability to give birth to progeny. This is seen especially in pregnant women (note that in human ancestors, pregnancy used to be much more frequent than today) who should pay extra attention to sources of pathogens in the environment to protect not only themselves but also their unborn babies (Tybur et al. 2011, 2013).

These findings have an interesting implication. There is an extensive line of evidence that pathogen disgust from potentially infectious stimuli has been transformed to the domain of

morality (Chapman et al. 2009; Pizarro et al. 2011). It is well supported in criminology statistics, that most of the crimes are committed by men (Rowe et al. 1995). It is possible, that higher propensity to feel disgusted is, on one hand, predisposing the person to suffer from some serious anxiety disorders, such as specific phobias (Page and Tan 2009; Thorpe and Salkovskis 1998; van Overveld et al. 2006) or OCD (McKay and Moretz 2009). On the other hand, it may serve as a protective factor helping individuals to obey social rules and avoid getting into trouble with the law. It could be suggested that people with biology education and a lower tendency to feel disgusted, such as men compared to women, are also more inclined to transgress moral codes of the society. This topic would require additional research.

Compliance with Ethical Standards

Conflict of Interest   The authors declare no conflict of interest.

Ethical Approval   All procedures performed in studies involving human participants were in accordance with the ethical standards of the appropriate institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent   Informed consent was obtained from all individual participants included in the study.

References

Addinsoft. (2017). *XLSTAT 2017: Data analysis and statistical solution for Microsoft Excel*. Paris: Addinsoft.

Arbuckle, J. L. (2016). *Amos (version 24.0) [computer program]*. Chicago: IBM SPSS.

Berger, U., & Anaki, D. (2014). Demographic influences on disgust: Evidence from a heterogeneous sample. *Personality and Individual Differences, 64*, 67–71.

Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science, 323*, 1222–1226.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation, 10*, 1–9.

Curtis, V. (2011). Why disgust matters. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 366*, 3478–3490.

Curtis, V., de Barra, M., & Aunger, R. (2011). Disgust as an adaptive system for disease avoidance behaviour. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 366*, 389–401.

Davey, G. C. (2011). Disgust: The disease-avoidance emotion and its dysfunctions. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 366*, 3453–3465.

Davey, G. C. L., & Marzillier, S. (2009). Disgust and animal phobias. In B. O. Olatunji & D. McKay (Eds.), *Disgust and its disorders: Theory, assessment, and treatment implications* (pp. 169–190). Washington, DC: American Psychological Association.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*, 169–200.

Faulkner, J., Schaller, M., Park, J. H., & Duncan, L. A. (2004). Evolved disease-avoidance mechanisms and contemporary xenophobic attitudes. *Group Processes & Intergroup Relations, 7*, 333–353.

Groth-Marnat, G. (2009). *Handbook of psychological assessment*. Hoboken: Wiley.

Haidt, J., McCauley, C., & Rozin, P. (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences, 16*, 701–713.

Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (Eds.). (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Lawrence Erlbaum Associates.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179–185.

IBM Corp. Released. (2013). *IBM SPSS statistics for windows, version 22.0*. Armonk: IBM Corp.

International Test Commission (2017). The ITC guidelines for translating and adapting test, 2nd ed. http://intestcom.org. Accessed 18 Sept 2017.

Kang, J. I., Kim, S. J., Cho, H. J., Jhung, K., Lee, S. Y., Lee, E., & An, S. K. (2012). Psychometric analysis of the Korean version of the disgust scale-revised. *Comprehensive Psychiatry, 53*, 648–655.

Kim, E. H., Ebesutani, C., Young, J., & Olatunji, B. O. (2013). Factor structure of the disgust scale–revised in an adolescent sample. *Assessment, 20*, 620–631.

Landová, E., Bakhshaliyeva, N., Janovcová, M., Peléšková, Š., Suleymanova, M., Polák, J., Guliev, A., & Frynta, D. (2018). Association between fear and beauty evaluation of snakes: Cross-cultural findings. *Frontiers in Psychology, 9*, 333.

LeBeau, R. T., Glenn, D., Liao, B., Wittchen, H. U., Beesdo-Baum, K., Ollendick, T., & Craske, M. G. (2010). Specific phobia: A review of DSM-IV specific phobia and preliminary recommendations for DSM-V. *Depression and Anxiety, 27*, 148–167.

McConnell, K. E., Strand, I. E., & Valdés, S. (1998). Testing temporal reliability and carry-over effect: The role of correlated responses in test-retest reliability studies. *Environmental and Resource Economics, 12*, 357–374.

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review, 15*, 28–50.

McKay, D., & Moretz, M. W. (2009). The intersection of disgust and contamination fear. In B. O. Olatunji & D. McKay (Eds.), *Disgust and its disorders: Theory, assessment, and treatment implications* (pp. 253–270). Washington, DC: American Psychological Association.

McLean, C. P., & Anderson, E. R. (2009). Brave men and timid women? A review of the gender differences in fear and anxiety. *Clinical Psychology Review, 29*, 496–505.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*, 396–402.

O'Connor, B. P. (n.d.). Programs for number of components and factors. Retrieved from https://people.ok.ubc.ca/brioconn/nfactors/nfactors.html.

Olatunji, B. O., & Sawchuk, C. N. (2005). Disgust: Characteristic features, social manifestations, and clinical implications. *Journal of Social and Clinical Psychology, 24,* 932–962.

Olatunji, B. O., Williams, N. L., Tolin, D. F., Abramowitz, J. S., Sawchuk, C. N., Lohr, J. M., & Elwood, L. S. (2007). The disgust scale: Item analysis, factor structure, and suggestions for refinement. *Psychological Assessment, 19,* 281–297.

Olatunji, B. O., Haidt, J., McKay, D., & David, B. (2008). Core, animal reminder, and contamination disgust: Three kinds of disgust with distinct personality, behavioral, physiological, and clinical correlates. *Journal of Research in Personality, 42,* 1243–1259.

Olatunji, B. O., Moretz, M. W., McKay, D., Bjorklund, F., de Jong, P. J., Haidt, J., Hursti, T. J., Imada, S., Koller, S., Mancini, F., Page, A. C., & Schienle, A. (2009). Confirming the three-factor structure of the disgust scale-revised in eight countries. *Journal of Cross-Cultural Psychology, 40,* 234–255.

Olatunji, B. O., Cisler, J., McKay, D., & Phillips, M. L. (2010). Is disgust associated with psychopathology? Emerging research in the anxiety disorders. *Psychiatry Research, 175,* 1–10.

Olatunji, B. O., Adams, T., Ciesielski, B., David, B., Sarawgi, S., & Broman-Fulks, J. (2012). The three domains of disgust scale: Factor structure, psychometric properties, and conceptual limitations. *Assessment, 19,* 205–225.

Olatunji, B. O., Ebesutani, C., Haidt, J., & Sawchuk, C. N. (2014). Specificity of disgust domains in the prediction of contamination anxiety and avoidance: A multimodal examination. *Behavior Therapy, 45,* 469–481.

Page, A. C., & Tan, B. J. (2009). Disgust and blood-injury-injection phobia. In B. O. Olatunji & D. McKay (Eds.), *Disgust and its disorders: Theory, assessment, and treatment implications* (pp. 191–209). Washington, DC: American Psychological Association.

Panter, A. T., Swygert, K. A., Grant Dahlstrom, W., & Tanaka, J. S. (1997). Factor analytic approaches to personality item-level data. *Journal of Personality Assessment, 68,* 561–589.

Pizarro, D., Inbar, Y., & Helion, C. (2011). On disgust and moral judgment. *Emotion Review, 3,* 267–268.

Polák, J., Sedláčková, K., Nácar, D., Landová, E., & Frynta, D. (2016). Fear the serpent: A psychometric study of snake phobia. *Psychiatry Research, 242,* 163–168.

Rowe, D. C., Vazsonyi, A. T., & Flannery, D. J. (1995). Sex differences in crime: Do means and within-sex variation have similar causes? *Journal of Research in Crime and Delinquency, 32,* 84–100.

Rozin, P., & Fallon, A. E. (1987). A perspective on disgust. *Psychological Review, 94,* 23–41.

Rozin, P., Haidt, J., & McCauley, C. (2000). Disgust. In M. Lewis & S. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 637–653). New York: The Guilford Press.

Rozin, P., Haidt, J., & McCauley, C. (2016). Disgust. In L. F. Barrett, M. Lewis, & M. Haviland-Jones (Eds.), *Handbook of emotions* (4th ed., pp. 815–834). New York: The Guilford Press.

Schienle, A. (2009). The functional neuroanatomy of disgust. In B. O. Olatunji & D. McKay (Eds.), *Disgust and its disorders: Theory, assessment, and treatment implications* (pp. 145–165). Washington, DC: American Psychological Association.

Schienle, A., Schäfer, A., Stark, R., Walter, B., & Vaitl, D. (2005). Gender differences in the processing of disgust-and-fear-inducing pictures: An fMRI study. *Neuroreport, 16,* 277–280.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Review, 34,* 1096–1109.

Schuirmann, D. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Pharmacodynamics, 15,* 657–680.

Stark, R., Walter, B., Schienle, A., & Vaitl, D. (2005). Psychophysiological correlates of disgust and disgust sensitivity. *Journal of Psychophysiology, 19,* 50–60.

Thorpe, S. J., & Salkovskis, P. M. (1998). Studies on the role of disgust in the acquisition and maintenance of specific phobias. *Behavior Research and Therapy, 36,* 877–893.

Tybur, J. M., Lieberman, D., & Griskevicius, V. (2009). Microbes, mating, and morality: Individual differences in three functional domains of disgust. *Journal of Personality and Social Psychology, 97,* 103–122.

Tybur, J. M., Bryan, A. D., Lieberman, D., Hooper, A. E. C., & Merriman, L. A. (2011). Sex differences and sex similarities in disgust sensitivity. *Personality and Individual Differences, 51,* 343–348.

Tybur, J. M., Lieberman, D., Kurzban, R., & DeScioli, P. (2013). Disgust: Evolved function and structure. *Psychological Review, 120,* 65–84.

van Overveld, W. J. M., de Jong, P. D., Peters, M. L., Cavanagh, K., & Davey, G. C. L. (2006). Disgust propensity and disgust sensitivity: Separate constructs that are differentially related to specific fears. *Personality and Individual Differences, 41,* 1241–1252.

van Overveld, M., de Jong, P. J., Peters, M. L., & Schouten, E. (2011). The disgust scale-R: A valid and reliable index to investigate separate disgust domains? *Personality and Individual Differences, 51,* 325–330.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432–442.