The DIsgust-RelaTed-Images (DIRTI) Database: Validation of a novel standardized set of disgust pictures

Anke Haberkamp,[1] Julia Anna Glombiewski,[1] Filipp Schmidt,[2] & Antonia Barke[1]

[1]Philipps-University Marburg, Germany
[2]Justus-Liebig-University Giessen, Germany

This is a post-peer-review, pre-copyedit version of an article published in ***Behaviour Research and Therapy.*** *The final authenticated version is available online at* http://dx.doi.org/10.1016/j.brat.2016.11.010

Abstract

Selecting appropriate stimuli is a major challenge of affective research. Although several standardized databases for affective pictures exist, none of them focus on discrete emotions such as disgust. Validated pictures inducing discrete emotions are still limited, and this presents a problem for researchers interested in studying different facets of disgust. In this paper, we introduce the DIsgust-RelaTed-Images (DIRTI) picture set. The set consists of 240 disgust-inducing pictures divided into six categories (*food*, *animals*, *body products*, *injuries/infections*, *death*, and *hygiene*). Additionally, we included 60 matched neutral pictures (10 per category). All pictures were rated by 200 participants on nine-point rating scales measuring *disgust*, *fear*, *valence,* and *arousal*. The present validation study covered a wide age range (18–75 years) with a balanced number of participants in each decade of life. For each picture, we provide separate ratings on the four scales for men and women. In addition to the original pictures, we also provide a luminance-matched version for experiments that require control of the physical properties of the pictures. The standardized DIRTI picture set allows researchers to chose from a wide set of disgust-inducing pictures and may enhance researchers' ability to draw comparisons between studies on disgust (download at http://dx.doi.org/10.5281/zenodo.167037).

Keywords: Disgust; Emotion; Picture Set; Affective visual stimuli

Introduction

Disgust has been counted among the basic emotions[1] since Darwin (1872), and like other basic emotions, elicits a stereotypical facial response (Ekman & Friesen, 1986), a characteristic physiological response (nausea), a specific behaviour (avoidance), and an emotional state (revulsion). Disgust-evoking objects tend to be those that are most likely contaminated by bacteria and viruses (Curtis, Aunger, & Rabie, 2004). Thus, disgust has a protective function for humans and animals. Several classification systems of disgust exist (Olatunji, Haidt, McKay, & David, 2008). The most widely accepted system was described by Rozin, Haidt, and McCauley (2000), who suggests a family of specialized forms of disgust, all of which are thought to have originated from the ancestral basic emotion. The authors distinguish between *core disgust* (i.e. (perceived) threat of oral incorporation elicited by rotten food, waste, body products and certain animals), *animal-reminder disgust,* which is said to remind us of our mortality (i.e. violations of the body, death), *interpersonal disgust* (i.e. contact with unknown persons, potentially carrying a disease), and *moral disgust* (i.e. moral violations; for a review see Chapman & Anderson, 2012). The latter form of disgust most likely also comprises other distinct emotions (i.e. anger and contempt; Olatunji et al., 2012).

In clinical research, disgust plays a major role in psychiatric disorders such as contamination-related obsessive-compulsive disorder (OCD) and several specific phobias (e.g. spider phobia or blood-injury-injection phobia; Cisler, Olatunji, Lohr, & Williams, 2009). In recent years, many clinical studies have investigated different aspects of disgust, using various methods to induce this emotion: for example, showing disgust-related videos (e.g. Sawchuk, Lohr, Lee, & Tolin, 1999) or pictures (e.g. Haberkamp & Schmidt, 2014), administering a bitter taste (Eskine, Kacinik, & Prinz, 2011), or using autobiographical recall (Fitzgerald et al., 2004).

It has been previously shown that visual material is effective in eliciting specific emotions (Lench, Flores, & Bench, 2011). When using pictorial stimuli, experimenters interested in inducing emotion in laboratory settings have to decide which pictures are suitable for that purpose. Researchers can use pictures from validated picture sets or can search for suitable pictures on the internet, and both methods have advantages and disadvantages; A number of validated, standardized sets of affective pictures with diverse content are available, including the International Affective Picture System (IAPS) (Lang, Bradley, & Cuthbert, 1999), the Nencki Affective Picture System (NAPS) (Marchewka, Żurawski, Jednoróg, & Grabowska, 2014), the Geneva Affective Picture Database (GAPED) (Dan-Glauser & Scherer, 2011), and the Emotional Picture System (EmoPicS) (Wessa et al., 2010).

Of the four general picture sets (IAPS, NAPS, GAPED, EmoPicS), the IAPS is the most widely used picture system (Marchewka et al., 2014). It is based on a dimensional approach to emotion (Osgood, Suci, & Tannenbaum, 1957; Wundt, 1896) and focuses on the dimensions of valence, arousal, and dominance. Additionally, a number of researchers have sought to classify the IAPS pictures into discrete categories of emotion (e.g. Barke, Stahl, & Kröner-Herwig, 2012; Davis et al., 1995; Libkuman, Otani, Kern, Viger, & Novak, 2007). However, despite its widespread use and established properties, the IAPS has certain disadvantages. First, contemporary clinical research mainly focuses on the study of distinct emotions (e.g., Barlow, 2000; Lawrence et al., 2007; Olatunji, Lohr, Sawchuk, & Tolin, 2006), which is not easily

[1] Emotions can be categorized as distinct or dimensional (cf. Barrett, & Wager, 2006). Here, we focus on the distinct emotion disgust since disgust plays a major role in different psychiatric disorders (cf. Cisler, Olatunji, Lohr, & Williams, 2009). Still, we are aware that a number of researchers prefer the categorization of emotion in terms of a dimensional approach.

reconciled with the dimensional approach of the IAPS. Second, even if IAPS pictures are assigned to distinct emotional categories, the number of pictures per category is rather limited. This is especially problematic because many contemporary research methods such as fMRI and EEG studies typically require a large number of pictures and may result in repetitive use of the same pictures. This may reduce the emotional induction effect (Marchewka et al., 2014) and introduce unwanted recognition effects. Third, the picture quality of IAPS pictures varies considerably. This may produce experimental artifacts, for example, if the picture quality of one emotional category is significantly poorer than the quality of a comparison category (Marchewka et al., 2014). Fourth, IAPS pictures also vary considerably with respect to basic features such as size, luminance, and complexity, which might also influence the emotional processing of these pictures (e.g. De Cesarei & Codispoti, 2006). Fifth, the original purpose of the IAPS pictures was to induce affective responses in non-clinical samples. As a consequence, some pictures show violent scenes (e.g. severely injured people, dead bodies), which creates two problems. First, using these pictures may raise ethical concerns. Secondly, because of their severity, these pictures tend to yield uniform emotional reactions regardless of participant characteristics; thus, patients – e.g. with anxiety disorders – as well as healthy controls are likely to respond in the same way, thereby obscuring potential differences between groups (Lissek, Pine, & Grillon, 2006).

Researchers have tried to overcome the physical disadvantages of the IAPS pictures by providing pictures in high-definition quality (see NAPS, GAPED and EmoPics). Furthermore, Riegel and colleagues (2016) provided ratings of basic emotions (happiness, anger, fear, sadness, disgust, and surprise), which facilitate the choice of appropriate images for researchers interested in these emotions and provide 51 images for disgust. However, since the Nencki Affective Picture System was not developed to address disgust specifically, the number of pictures per emotion is still limited, and is even smaller for categories of disgust (e.g. food, animals, hygiene). Consequently, for any researcher specifically interested in studying different categories of disgust, the existing sets (NAPS, GAPED, and EmoPics) fall short with respect to number of pictures and some (GAPED and EmoPIcs) have not been validated for discrete emotions.

In this situation, many researchers resort to collecting pictures from the internet and compile their own custom-made stimulus set (e.g. Buodo, Peyk, Junghöfer, Palomba, & Rockstroh, 2007; Haberkamp & Schmidt, 2014). However, this approach is time-consuming and may cause problems with copyright legislation. More importantly, each researcher might collect pictures relying on his or her own conception of the particular emotion, which may well differ from the participants' views (Barke et al., 2012). Although the stimulus material is often rated in the course of the experiment, these ratings are post hoc and cannot guide picture selection. Finally, the use of custom-made stimulus sets necessarily jeopardizes comparability across studies.

To resolve these issues, we developed a picture set to study the emotion of disgust. To our knowledge, this is the first validated picture set for the induction of disgust. We decided to include pictures for those disgust categories that fulfill the following criteria: each category should (1) primarily address the emotion of disgust, (2) be unambiguously related to its content (e.g., not include complex interpersonal situations), and (3) play a major role in psychiatric disorders (i.e. OCD, specific phobias). These criteria were fulfilled by six disgust categories: (a) *food* (e.g. spoiled food), (b) *animals*[2] (e.g. worms, cockroaches), (c) *body products* (e.g. feces), (d)

[2] Note that spiders and snakes were excluded from this category because rated pictures of those stimuli are already covered by the GAPED database.

injuries/infections (e.g. skin rashes, lesions), (e) *death* (e.g. animal cadavers, bones), and (f) *hygiene* (e.g. dirty bathrooms). We aimed to cover a broad range of disgust intensity in each category, from mildly to moderately to highly disgusting pictures. However, we avoided extremely disgust-provoking pictures for ethical and experimental reasons (e.g. provoking uniform responses for these pictures across experimental and control groups; Lissek et al., 2006).

Methods

The study was approved by the Ethical Committee of the Faculty of Psychology (Philipps-University Marburg).

Selection of Pictures. The majority of pictures were collected from the internet (www.flickr.com), and some additional photographs were taken by one co-author (AB) and two graduate students. All pictures are copyright-free and covered by creative commons licences (i.e., there are no restrictions on these pictures with regard to copying, editing, and distribution). Additionally, the photographers and the individuals in the photographs gave written informed consent for the use of the pictures for scientific purposes. A large pool of pictures was chosen according to their content (i.e., whether they represented one of the six disgust categories of *food, animals, body products, injuries/infections, death, or hygiene*) and picture quality (i.e. sharpness, noise, luminance, contrast, distortion etc.). All pictures were in landscape format, and pictures with visible commercial logotypes were removed. Large written words were removed to avoid attentional effects and to make the picture sets less culture-specific. For ethical reasons, the category *death* only contained dead animals (rather than people).

The initial picture pool preselected by the authors resulted in 356 potentially disgust-provoking pictures covering the six categories. These pictures were edited to achieve a uniform size (1024 x 768 pixels) and to adjust the picture parameters to ensure that they had a reasonably even color tone, contrast, and lighting. To select the final pictures for the database, 10 participants (five females, five males) rated all 356 pictures in a pilot study according to their content (i.e. 'Which category is presented in this picture?') and their appropriateness for a disgust database (ranging from 0 = 'entirely inappropriate' to 4 = 'entirely appropriate'). Additionally, participants were encouraged to comment on the presented pictures. Each picture was announced by an audible signal and was presented for 12 seconds. A total of 240 disgust pictures (40 per category) were selected based on these ratings. To cover a broad range of stimuli, we excluded pictures that were very similar to one another. Subsequently, 60 neutral pictures were assembled (10 per category). In each category, we aimed to match the content of the neutral pictures to the content of the disgust-related pictures. For example, we collected sleeping animals as neutral control pictures for the dead animals in the category of death. For the category of food, we collected pictures of fresh fruit; for the category of body products, we collected pictures of clean bathrooms, towels etc.; for the category of injuries/infections, we collected pictures of unharmed body parts, and for the category of hygiene, we collected pictures of clean sanitary articles.

We also report the following physical properties for each picture: grayscale luminance and contrast, mean channel values in CIE 1976 L*a*b color space, and JPEG file size. Grayscale luminance and contrast were obtained by converting images to grayscale and computing the mean (luminance) and standard deviation (contrast) of all pixel values. Mean channel values in CIE 1976 L*a*b color space were obtained by converting RGB values to color space values and computing the mean of each channel. As CIE 1976 L*a*b is a color-opponent space, it approximates characteristics of the human visual system with lightness dimension L* [0...100] and color-opponent dimensions a* (green-red) and b* (yellow-blue) [−127...+128]. Finally, we report JPEG compression file size (with maximal compression quality setting), as in detailed color images this value can

correlate substantially with subjective measures of image complexity (Donderi, 2006). Physical properties for each image are listed in the supplementary material. Additionally, we provide images that were matched with respect to their average luminance by matching them in the lightness dimension of CIE 1976 L*a*b color space. Note that the ratings were based exclusively on the original images. For researchers who wish to use the original pictures but also want to control for luminance, we identified 10 subsets of the original images, each of which contains images of comparable luminance across the 7 categories (food, animals, body products, injuries/infections, death, hygiene, and neutral; cf. Supplementary material).

Participants. A total of 200 volunteers from the general population participated in the study (102 women, 98 men, mean age = 43.5 years, *SD* = 14.9, age range = 18–75 years). We aimed to cover a broad age range in our sample. Accordingly, participants were age-stratified, and once enough participants in one age decade (e.g. ranging from 30–39 years) were recruited, we stopped recruiting participants in that age range. Thus, we obtained a balanced number of participants across age groups (age group 18–29 years: *n* = 44; age group 30–39 years: *n* = 44; age group 40–49 years: *n* = 38; age group 50–59 years: *n* = 38; age group 60–75 years: *n* = 36). Participants were recruited via newspapers and bulletin boards or were approached in person by one of two graduate student research assistants. No participants were receiving psychological or psychiatric treatment, and all had stated that they were healthy. All participants had normal or corrected-to-normal visual acuity. Twenty-six participants received payment of €10, while 174 participants took part without financial compensation. All participants gave informed consent and were treated in accordance with the ethical guidelines of the American Psychological Association.

Apparatus, picture presentation and rating scales. Pictures were presented on a laptop with a 15-inch display and a resolution of 1366 x 768 pixels. Participants were seated in a quiet room in front of the laptop at a viewing distance of approximately 70 cm. Before the experimental session started, participants were given details about the contents of the pictures. Participants were informed that if they felt any discomfort during the session, they should report it to the investigator in order to stop the rating session. The pictures were presented using a custom-made computer program (programmed in Visual Basic 6.0, Microsoft), which displayed each picture individually in the top part of the screen, as well as four rating scales below the picture for valence, arousal, disgust, and fear (see below). Each participant viewed all 300 pictures (240 disgust pictures plus 60 neutral pictures). The order of the picture presentation was fully randomised. Each picture remained on the screen until the participant finished his or her ratings. As noted above, participants rated each picture with respect to *valence*, *arousal*, *disgust*, and *fear*. Thus, the pictures were rated on a dimensional level (valence and arousal) as well as a categorical level (disgust and fear). The valence scale ranged from 1 = *very negative* to 9 = *very positive* with 5 = *neutral*. The arousal, disgust, and fear scales ranged from 1 = *none* to 9 = *very strong*; thus they were coded in such a way that higher scores reflected higher arousal, stronger disgust, and stronger fear, respectively. Participants rated the pictures by ticking the appropriate option buttons. In addition, they completed a German questionnaire that measures disgust sensitivity, developed by Schienle, Walter, Stark, and Vaitl (2002) (FEE 'Fragebogen zur Erfassung der Ekelempfindlichkeit'). The experiment lasted approximately 1.5 hours. Participants were allowed to take breaks at any time.

Data treatment and statistical methods. The program recorded a log file for each participant. The log files were imported into Excel by means of a custom macro. No data were excluded.

For each individual picture, the mean ratings and standard deviations for disgust, fear, valence, and arousal were calculated for men and women, as well as for men and

women combined. In order to characterize the picture categories, the average ratings for individual pictures were then used to compute a grand mean for disgust, fear, valence, and arousal for each category. We also calculated the Pearson correlations between the disgust ratings and the fear ratings, valence ratings, and arousal ratings of the pictures.

At the participant level, we calculated the correlation between disgust ratings and age. To test for sex differences in ratings for each of the categories, we calculated four 2x6 mixed design ANOVAS with sex as the between participants factor and category (disgust, fear, valence, or arousal) as the within participants factor. As measures of effect size, Cohen's *d* and η^2 are reported as appropriate (cf. Levine & Hullet, 2002). Note that according to Cohen (1988) an effect size (η^2) of 0.01 reflects a small effect, of 0.059 a medium effect, and of 0.138 a large effect. In addition, we investigated the internal consistency of the data using a split-half reliability estimate (Wierzba et al., 2015).

Results

Ratings. Ratings for each individual picture were calculated by computing the mean value and the standard deviation for disgust, fear, valence, and arousal. The standard deviation here represents the amount of variation among the individuals rating each picture. These data may help researchers in selecting stimulus material and can be downloaded in the upplementary material. Example pictures with low, medium, and high scores for disgust ratings as well as neutral examples in each category are displayed in Figure 1.

To provide summary data, we also calculated the means for each category by computing the mean value and standard deviation for disgust, fear, valence, and arousal across the different pictures belonging to the category (Tables 1 and 2). Here, the standard deviation represents the amount of variation among the pictures belonging to one category. For the ratings of the different qualities (disgust, fear, valence and arousal) for each category, see Figure 2.

Figure 1. Pictures scoring lowest (1st column), medium (2nd column), and highest (3rd column) in disgust ratings next to a neutral example (4th column) for each category from the DIRTI pictures set. (For a list of photographers of the respective pictures, see the Supplementary material)

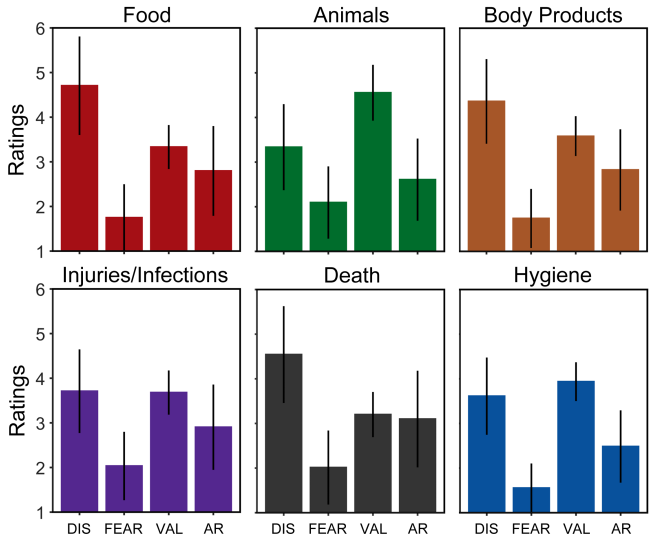

Figure 2. Means and standard deviations for disgust, fear, valence, and arousal ratings in each category. Disgust, fear, and arousal were rated from 1 = *none* to 9 = *very strong*; note that valence is inversely coded ranging from 1 = *very negative* to 9 = *very positive*.

Table 1. Mean, standard deviation, minimum and maximum for disgust and fear ratings per category.

	Disgust				Fear			
	M	*SD*	Min	Max	*M*	*SD*	Min	Max
Body products	4.36	1.20	1.62	6.92	1.74	0.27	1.28	2.33
Death	4.54	0.89	2.29	6.18	2.01	0.23	1.53	2.67
Food	4.70	0.62	3.37	5.84	1.75	0.18	1.41	2.10
Animals	3.33	0.63	2.21	4.62	2.09	0.35	1.42	3.04
Injuries/infections	3.71	0.98	1.81	5.49	2.03	0.33	1.27	2.65
Hygiene	3.61	0.98	1.31	6.44	1.55	0.24	1.14	2.37
Disgust Total	4.04	1.04	1.31	6.92	1.86	0.33	1.14	3.04
Neutral Total	1.12	0.09	1.01	1.39	1.13	0.13	1.01	1.74

Table 2. Mean, standard deviation, minimum and maximum for valence and arousal ratings per category.

	Valence				Arousal			
	M	*SD*	Min	Max	*M*	*SD*	Min	Max
Body products	3.58	0.74	2.07	5.73	2.82	0.59	1.78	4.30
Death	3.20	0.52	2.44	4.74	3.10	0.43	2.10	4.02
Food	3.33	0.37	2.69	4.21	2.80	0.32	2.14	3.47
Animals	4.55	0.56	3.61	5.52	2.61	0.37	1.89	3.40
Injuries/infections	3.68	0.53	2.79	4.85	2.90	0.51	1.62	3.83
Hygiene	3.94	0.74	2.20	6.47	2.48	0.51	1.32	4.09
Disgust Total	3.71	0.74	2.07	6.47	2.79	0.50	1.32	4.30
Neutral Total	7.30	0.74	5.88	8.38	1.39	0.17	1.14	1.86

Correlations between Disgust, Fear, Valence, and Arousal. For the correlations among the ratings across all pictures see Table 3; for scatter plots for each category see Figure 3. In all categories, there were linear relationships of similar strength between valence and disgust and arousal and disgust: the higher the disgust, the lower the valence and the higher the arousal. There was also a positive relationship between disgust and fear. It shouldbe noted that in absolute terms, the pictures evoked much more disgust than fear (maximum fear rating for a picture = 3.04 for an animal picture; maximum disgust rating for a picture = 6.92 for a body product picture).

Matched neutral pictures. For each category, we provide 10 neutral pictures matched for content matter. In each category, the neutral pictures were regarded as significantly less disgusting (all *t*s > 17.0), less fear-evoking (all *t*s > 7.0), less arousing (all *t*s > 9.0), and possessing higher valence (all *t*s > 29.0) for all categories, with $p < .000001$ for all rated variables.

Table 3. Correlations between the ratings of disgust, fear, valence and arousal for the DIRTI pictures. All reported correlations are significant after Bonferroni correction ($p < .008$)

	M	*SD*	Disgust	Fear	Valence
Disgust	4.0	1.77			
Fear	1.9	1.34	.64		
Valence	3.7	0.81	-.63	-.46	
Arousal	2.8	1.78	.69	.77	-.46

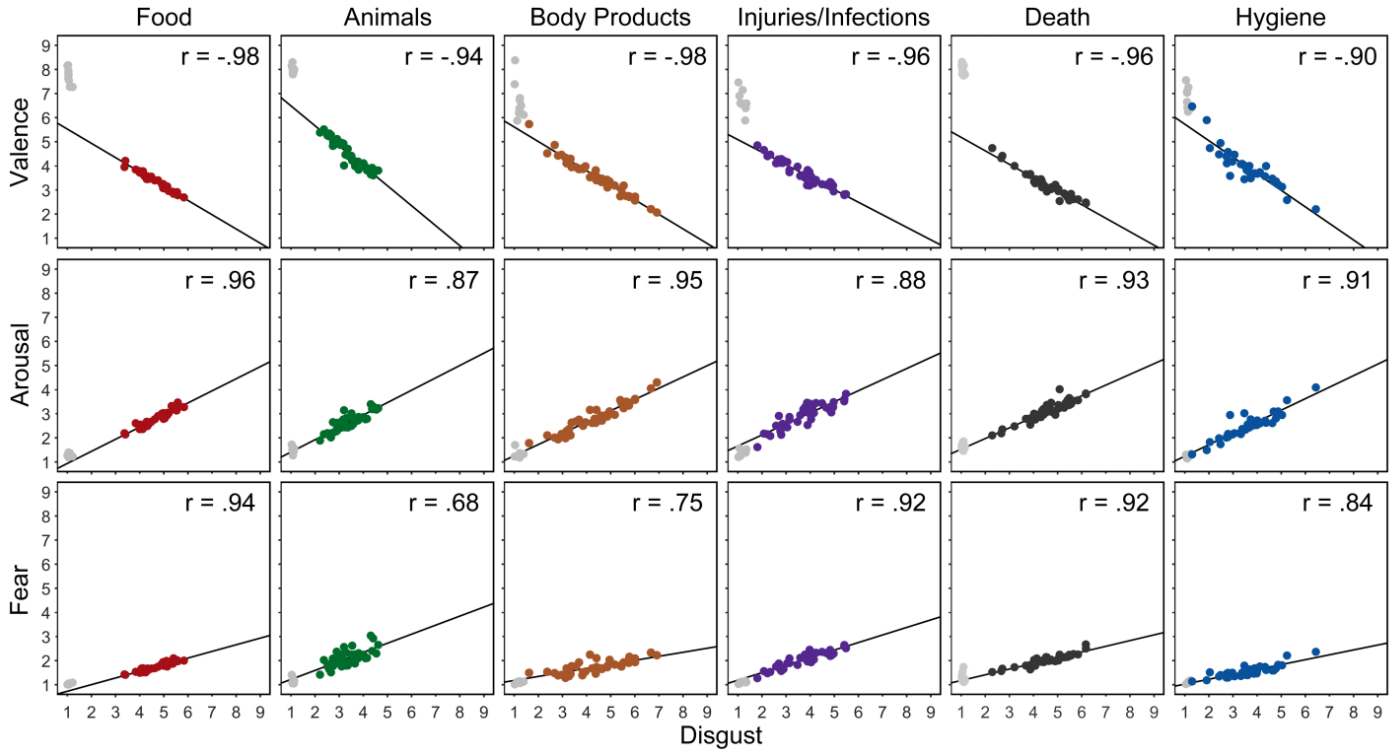

Figure 3. Correlations between fear, arousal, valence, and disgust for each category. Each dot represents one rated picture. Colored dots represent disgust-related pictures. Grey dots represent matched neutral pictures for each category; they were not included in the calculations of the correlations, but are displayed here for the purpose of comparison. Disgust, fear, and arousal were rated from 1 = none to 9 = very strong; note that valence is inversely coded ranging from 1 = very negative to 9 = very positive.

All statistical values are available in the supplementary material.

Age. Age was unrelated to the disgust ratings as a whole (r = −.02) as well as the disgust ratings for the individual categories (food r = .01, animals r = −.06, body products r = −.02, injuries/infections r = −.14, death r = .05, hygiene r = .04).

Sex differences. Overall, the ratings of the disgust pictures were highly correlated between men and women for disgust (r = .96), valence (r = .96), arousal (r = .91), and fear (r = .90), all p < .001. These high correlations indicated that the relative ranking of the pictures for men and women was very similar. Men's and women's ratings for the neutral pictures were also correlated for disgust (r = .56), valence (r = .96), arousal (r = .68) and fear (r = .84), all p < .001. However, the actual value of the ratings for individual pictures may still differ for men and women despite high overall correlations. For this reason, we also examined sex differences for ratings of each picture (Supplementary material). Overall, we found significant (p < .05) sex differences in disgust ratings for 126 pictures and in valence ratings for 92 pictures, but sex differences on arousal ratings were only found for four pictures and differences on fear ratings were only found for three pictures. The effect sizes of the significant differences ranged from small to medium as expressed by Cohen's d: for disgust, the differences were medium for 17 pictures (i.e., d ≥ |0.50|) and small for 109 pictures (i.e., |0.28| ≥ d ≤ |0.49|); for valence, differences were medium for 12 pictures and small for 80 pictures; for arousal, the difference was medium for one picture and small for three pictures; and for fear, the differences were small for all three pictures.

On an aggregate level, we calculated four 2x6 mixed design ANOVAs with the factors of sex (men, women) and picture category (food, animals, body products, injuries/infections, death, hygiene) for the outcome variables of disgust, valence, arousal and fear.

Disgust. For the disgust ratings, the 2x6 ANOVA with the factors of sex and picture category showed main effects for sex ($F(1,198) = 7.14$, $p = .008$, $\eta^2 = 0.027$) and picture category ($F(5,990) = 88.21$, $p < .0001$, $\eta^2 = 0.066$), and an interaction between sex and picture category ($F(5,990) = 4.12$, $p = .001$, $\eta^2 = 0.003$). Bonferroni-corrected post hoc tests among the picture categories (all category combinations were tested and we corrected for the number of tests) showed significant differences between all categories except; "death-body products", "death-food", and "hygiene-injuries/infections". Planned contrasts (Bonferroni-corrected threshold $p < .0083$) showed that women rated the picture categories of food ($t(198) = 2.69$, $p < .008$, $d = 0.38$), body products ($t(198) = 2.90$, $p < .008$, $d = 0.41$), death ($t(198) = 3.05$, $p < .008$, $d = 0.43$), and hygiene ($t(198) = 3.12$, $p < .008$, $d = 0.44$) as more disgusting than did men. No sex differences were found for the categories of animals and injuries/infections (Figure 4, upper left panel).

Fear. The 2x6 ANOVA with the factors of sex and picture category showed a main effect for picture category ($F(5, 990) = 32.44$, $p < .0001$, $\eta^2 = 0.018$), but no effect for sex ($F(1,198) = 0.02$, $p = .88$, $\eta^2 = 0.0001$) and no interaction between sex and picture category ($F(5, 990) = 1.63$, $p = 0.15$, $\eta^2 = 0.0009$). Bonferroni-corrected post hoc tests for category differences showed differences between all picture categories except "food-body products", "animals-death", "animals-injuries/infections", "injuries/infections-death" (Figure 4, upper right panel; as above, all category combinations were tested and we corrected for the number of tests).

Valence. For valence, the 2x6 ANOVA with the factors of sex and picture category yielded main effects for sex ($F(1, 198) = 5.57$, $p = .02$, $\eta^2 = 0.016$) and picture category ($F(5,990) = 144.20$, $p < .0001$, $\eta^2 = 0.174$) and an interaction between sex and picture category ($F(5,990) = 3.29$, $p = .006$, $\eta^2 = 0.004$). Bonferroni-corrected post hoc tests for category (all category combinations were tested and we corrected for the number of tests) showed differences between all picture categories except "body products-injuries/infections" and "death-food". Planned contrasts (Bonferroni-corrected threshold $p < .0083$) indicated that women rated the categories of body products ($t(198) = 2.68$, $p < .008$, $d = 0.38$) and death ($t(198) = 3.63$, $p < .008$, $d = 0.51$) as more negative than did men (Figure 4, lower left panel).

Arousal. With regard to arousal, the 2x6 ANOVA with the factors of sex and picture category showed a main effect for picture category ($F(5,990) = 26.30$, $p < .0001$, $\eta^2 = 0.011$), but not for sex ($F(1,198) = 0.16$, $p = .69$, $\eta^2 = 0.0007$). There was a significant sex x picture category interaction ($F(5,990) = 2.64$, $p = .02$, $\eta^2 = 0.001$). Bonferroni-corrected post hoc tests for category (all category combinations were tested and we corrected for the number of tests) indicated significant differences between all categories except "body products-food", "body products-injuries/infections", "food-injuries/infections" and "hygiene-animals". Bonferroni-corrected planned contrasts for sex differences for the individual picture categories showed no significant differences (Figure 4, lower right panel).

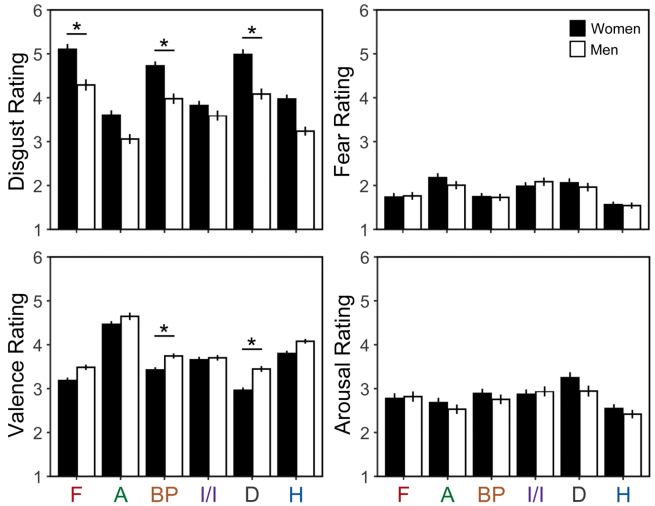

Figure 4. Disgust, fear, valence, and arousal ratings of men and women for the picture categories. Error bars show standard errors. Disgust, fear, and arousal were rated from 1 = *none* to 9 = *very strong*; note that valence is inversely coded ranging from 1 = *very negative* to 9 = *very positive*. Categories: F = Food; A = Animals; BP = Body Products; I/I = Infections/Injury; D = Death; H = Hygiene.

Asterisks mark post hoc tests significant at $p < .0083$ (Bonferroni corrected threshold). Disgust-Ratings: F: $t(198) = 2.69$, $d = 0.38$; BP: $t(198) = 2.89$, $d = 0.41$; D: $t(198) = 3.05$, $d = 0.43$; H: $t(198) = 3.12$, $d = 0.44$; Valence-Ratings: BP: $t(198) = 2.68$, $d = 0.38$; D: $t(198) = 3.63$, $d = 0.51$.

Disgust sensitivity. Participants also completed a questionnaire, the FEE, measuring disgust sensitivity. Overall, women (mean item score 2.01 ± 0.48) reported higher disgust sensitivity than men (1.68 ± 0.56) ($t(198) = 4.50$, $p < .0001$, $d = 0.64$). The ratings of the disgust-evoking pictures correlated highly with disgust sensitivity as measured by the FEE ($r = .60$).

Reliability. We estimated the internal consistency of participants' ratings by calculating split-half reliability scores (Wierzba et al., 2015). To this end, participants were numbered according to their order of participation and were split into two groups (i.e., odd vs. even participant numbers). Next, we calculated the average ratings for disgust, fear, valence, and arousal separately for each image and within each participant group. Finally, we calculated Pearson correlations among these average ratings for the two participant groups (1) across all disgust categories and (2) within each category. All correlations were significant ($p < .001$), and Spearman-Brown corrected reliability scores were high (1) across all disgust images ($r = 0.99$ for disgust; $r = 0.97$ for fear; $r = 1.00$ for valence; $r = 0.99$ for arousal) but also (2) within each category (all r's > 0.97 for disgust; all r's > 0.87 for fear; all r's > 0.96 for valence; all r's > 0.92 for arousal). The magnitudes of the split-half reliability scores are comparable to those previously reported for other standardized image sets (e.g., Lang, Bradley, & Cuthbert, 1999).

Discussion

With the present DIRTI picture set, we provide a comprehensive set of 240 static, disgust-related pictures from six categories (food, animals, body products, injuries/infections, death, and hygiene). In addition, we collected 60 neutral pictures, 10 for each category. All pictures were rated with respect to the categorical emotions of disgust and fear, and were also rated on the dimensions of valence and arousal. Pictures and ratings by men and women are available to researchers via LINK TO SUPPLEMENTARY MATERIALTO BE INSERTED HERE.

Overall, the 240 disgust pictures evoked medium to strong disgust. Although disgust ratings were moderately correlated with ratings of fear, the fear ratings were generally low (cf. Table 1). Thus, a disgust picture from this set generally induces moderate to high levels of disgust but only low levels of fear. In general, the disgust pictures were also rated as negatively valenced and moderately arousing. The disgust ratings and the dimensions of valence and arousal are related in the expected manner: the higher the disgust rating for a picture, the lower its valence and the higher the associated arousal. This was true for the full picture set and also held within each category. In each category, we aimed to match the content of the neutral pictures to the subject matter of the disgust-related pictures. For example, in the category of death, we collected sleeping animals as control pictures, and for hygiene, we used pictures of clean rather than soiled bathrooms. As expected, neutral pictures were rated as evoking no disgust, fear, or arousal, and as being more neutrally valenced, permitting a choice of validated neutral stimuli matched in content and picture parameters. Also, the magnitudes of the split-half reliability scores were high across all disgust images as well as within each category.

Among the disgust pictures, we strove to include pictures evoking different levels of disgust to allow researchers to choose the appropriate strength of disgust for a given study. This is important because stimuli that are too strong may eliminate differences between participants that may be of interest to a particular study. Lissek and colleagues (2006) refer to the psychological concept of the strong situation (Ickes, 1982; Mischel, 1977; Monson & Snyder, 1977) in which highly intense stimuli yield uniform reactions and

expectancies irrespective of person characteristics. By contrast, in a weak situation, stimuli are of lower salience, diminishing the situational influence and allowing person characteristics to determine behaviour.

Our results show that disgust ratings of men and women were highly correlated, indicating that the ranking of pictures is very similar for both sexes. Nevertheless, we found differences in the absolute level of disgust and valence reported in around one third of pictures, but almost no sex differences were found for fear and arousal. We would therefore advise researchers to take these findings into account and to consult the tables reporting the mean ratings for men and women separately when choosing stimuli according to disgust and valence ratings.

We found that the sex differences were not equally distributed across all categories, but that women rated the categories of food, body products, and death as more disgusting than did men. Sex differences in disgust ratings have been previously reported described: women tend to report feelings of disgust more often and rate this emotion as stronger compared to men (Bradley, Codispoti, Sabatinelli, & Lang, 2001). Women also showed more disgust sensitivity in the questionnaire in the present study as well as in the literature (Olatunji, Sawchuk, Arrindell, & Lohr, 2005).

The question of whether disgust is influenced by an individual's age has remained contentious. To address this question and to ensure that the collected ratings would be useful across a wide range of age groups, we recruited participants in each age decade from young adulthood to middle age and beyond. Age was unrelated to the disgust ratings as a whole and in each individual category. This is noteworthy, since two contradictory theories exist regarding the putative relationship between age and disgust sensitivity. Proponents of the Terror Management Theory (Greenberg, Solomon, & Pyszczynski, 1997) argue that existential anxiety arises every time we are reminded of our own mortality by death-related disgust stimuli. The disgust response to these stimuli helps to avoid this existential anxiety. According to this theory, disgust sensitivity should increase in older age as the time of death approaches. By contrast, disgust sensitivity could be thought to decline with age, reflecting a habituation process to disgust-related stimuli over the individual lifetime (Fessler & Navarrete, 2005). Our findings do not support either of these theories. In addition, we collected data on the self-reported disgust sensitivity of the participants and again found no correlation between disgust sensitivity and age. This is in line with the results of Schienle and colleagues (2002), who found no correlation between age and disgust, and Curtis and colleagues (2004), who found only a small downward trend for disgust sensitivity with increasing age. Our results suggest that the DIRTI picture set is age-independent and selection of stimuli can proceed without reference to the participants' age.

In some experimental designs, response times have been shown to depend on purely physical properties of an image such as luminance and contrast (Plainis & Murray, 2000), although both may be less influential in speeded categorisation of natural images (Macé, Delorme, Richard, & Fabre-Thorpe, 2010). Therefore, we decided to provide a luminance-matched version of the DIRTI picture set for those researchers who aim to control for low-level differences between different pictures. However, it is important to note that only the original DIRTI picture set was rated by the participants, and luminance matching subtly changes the appearance of the original pictures. Thus, although we expect that the ratings of the luminance-matched pictures will not differ substantially from the ratings of the original pictures, we recommend using the original picture set if low-level characteristics are not crucial. If researchers employ the luminance-matched pictures, we advise collecting additional ratings in the course of the study.

The present study is the first to produce a large set of disgust-related stimuli rated by a sizeable sample of men and women across all

age groups. A limitation of our study is that the luminance-matched version of the DIRTI picture set was not rated by the participants. This reflects our judgment that disgust stimuli are best used in their original form, but we cannot rule out the possibility that the disgust-inducing effect might vary between the two versions of the picture set. In addition, our participants were a sample from the general population; they were asked about their health, but were not screened for mental disorders. Therefore, we cannot rule out the possibility that some participants may have suffered from an undisclosed mental disorder. However, in this respect our sample was comparable to other samples in similar studies, and the broad age range and use of a general population sample rather than a student sample represent strengths of the current study.

DIRTI may also have clinical applications, and might be particularly useful for clinicians who work with individuals suffering from OCD or disgust-related specific phobias. Specifically, patients' DIRTI ratings could be used to design exposures for fear of contamination with different levels of difficulty. DIRTI could also serve as material for exposures or might help in evaluating the outcome of exposures. For individuals with OCD, we would expect fear ratings to be higher than for volunteers from the general population and to be correlated with disgust. Clearly, further research on this subject is needed. A study evaluating DIRTI for clinical populations is currently in preparation.

In summary, our picture set DIRTI (DIsgust-RelaTed Images) and the accompanying ratings are available to researchers for (noncommercial) scientific purposes. It was compiled in the hope that it will prove useful to researchers when selecting stimuli and will facilitate comparisons among studies on disgust.

Acknowledgments

We would like to thank our colleagues for their support and helpful discussions, the QSL-commission of the University of Marburg for financial support, and Angelika Gosch and Melissa Muster for data collection.

References

Barke, A., Stahl, J., & Kröner-Herwig, B. (2012). Identifying a subset of fear-evoking pictures from the IAPS on the basis of dimensional and categorical ratings for a German sample. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(1), 565–572.

Barlow, D. H. (2000). Unraveling the mysteries of anxiety and its disorders from the perspective of emotion theory. *American Psychologist*, *55*(11), 1247.

Barrett, L. F., & Wager, T. D. (2006). The structure of emotion evidence from neuroimaging studies. *Current Directions in Psychological Science*, *15*(2), 79-83.

Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: sex differences in picture processing. *Emotion*, *1*(3), 300.

Buodo, G., Peyk, P., Junghöfer, M., Palomba, D., & Rockstroh, B. (2007). Electromagnetic indication of hypervigilant responses to emotional stimuli in blood-injection-injury fear. *Neuroscience Letters*, *424*(2), 100–105.

Chapman, H. A., & Anderson, A. K. (2012). Understanding disgust. *Annals of the New York Academy of Sciences*, *1251*(1), 62–76.

Cisler, J. M., Olatunji, B. O., Lohr, J. M., & Williams, N. L. (2009). Attentional bias differences between fear and disgust: Implications for the role of disgust in disgust-related anxiety disorders. *Cognition and Emotion*, *23*(4), 675–687.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* Hillsdale, NJ: Lawrence Erlbaum Associates.

Curtis, V., Aunger, R., & Rabie, T. (2004). Evidence that disgust evolved to protect from risk of disease. *Proceedings of the Royal Society of London B: Biological Sciences*, *271*(Suppl 4), S131–S133.

Dan-Glauser, E. S., & Scherer, K. R. (2011). The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. *Behavior Research Methods*, *43*(2), 468–477.

Darwin, C. (1872). 1965. *The expression of the emotions in man and animals*. London, UK: John Marry.

Davis, W. J., Rahman, M. A., Smith, L. J., Burns, A., Senecal, L., McArthur, D., ... & Wagner, W. (1995). Properties of human affect induced by static color slides (IAPS): dimensional, categorical and electromyographic analysis. *Biological Psychology*, *41*(3), 229–253.

De Cesarei, A., & Codispoti, M. (2006). When does size not matter? Effects of stimulus size on affective modulation. *Psychophysiology*, *43*(2), 207–215.

Donderi, D. C. (2006). An information theory analysis of visual complexity and

dissimilarity. *Perception*, *35*, 823.

Ekman, P., & Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. *Motivation and Emotion*, *10*(2), 159–168.

Eskine, K. J., Kacinik, N. A., & Prinz, J. J. (2011). A bad taste in the mouth gustatory disgust influences moral judgment. *Psychological Science*, *22*(3), 295–299.

Fessler, D. M., & Navarrete, C. D. (2005). The effect of age on death disgust: challenges to terror management perspectives. *Evolutionary Psychology*, 3, 279–296.

Fitzgerald, D. A., Posse, S., Moore, G. J., Tancer, M. E., Nathan, P. J., & Phan, K. L. (2004). Neural correlates of internally-generated disgust via autobiographical recall: a functional magnetic resonance imaging investigation. *Neuroscience Letters*, *370*(2), 91–96.

Greenberg, J., Solomon, S., & Pyszczynski, T. (1997). Terror management theory of self-esteem and cultural worldviews: Empirical assessments and conceptual refinements. *Advances in Experimental Social Psychology,* 29, 61–139.

Haberkamp, A., & Schmidt, T. (2014). Enhanced visuomotor processing of phobic images in blood-injury-injection fear. *Journal of Anxiety Disorders*, *28*(3), 291–300.

Ickes, W. (1982). A basic paradigm for the study of personality, roles, and social behavior. In *Personality, roles, and social behavior* (pp. 305–341). New York: Springer.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). International affective picture system (IAPS): Instruction manual and affective ratings. *The center for research in psychophysiology, University of Florida*.

Lawrence, N. S., An, S. K., Mataix-Cols, D., Ruths, F., Speckens, A., & Phillips, M. L. (2007). Neural responses to facial expressions of disgust but not fear are modulated by washing symptoms in OCD. *Biological psychiatry*, *61*(9), 1072-1080.

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitations. *Psychological Bulletin*, *137*(5), 834.

Levine, T. R., & Hullett, C. R. (2002). Eta squared, partial eta squared, and misreporting of effect size in communication research. *Human Communication Research*, *28*(4), 612-625.

Libkuman, T. M., Otani, H., Kern, R., Viger, S. G., & Novak, N. (2007). Multidimensional normative ratings for the international affective picture system. *Behavior Research Methods*, *39*(2), 326–334.

Lissek, S., Pine, D. S., & Grillon, C. (2006). The strong situation: A potential impediment to studying the psychobiology and pharmacology of anxiety disorders. *Biological Psychology*, *72*(3), 265–270.

Macé, M. J. M., Delorme, A., Richard, G., & Fabre-Thorpe, M. (2010). Spotting animals in natural scenes: efficiency of humans and monkeys at very low contrasts. *Animal Cognition*, *13*(3), 405–418.

Marchewka, A., Żurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, *46*(2), 596–610.

Mischel, W. (1977). The interaction of person and situation. *Personality at the crossroads: Current issues in interactional psychology*, *333*, 352.

Monson, T. C., & Snyder, M. (1977). Actors, observers, and the attribution process: Toward a reconceptualization. *Journal of Experimental Social Psychology*, *13*(1), 89–111.

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*(2), 205–225.

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*(5), 1243–1259.

Olatunji, B. O., Lohr, J. M., Sawchuk, C. N., & Tolin, D. F. (2007). Multimodal assessment of disgust in contamination-related obsessive-compulsive disorder. *Behaviour research and therapy*, *45*(2), 263-276.

Olatunji, B. O., Sawchuk, C. N., Arrindell, W. A., & Lohr, J. M. (2005). Disgust sensitivity as a mediator of the sex differences in contamination fears. *Personality and Individual Differences*, *38*(3), 713–722.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957) *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.

Plainis, S., & Murray, I. J. (2000). Neurophysiological interpretation of human visual reaction times: effect of contrast, spatial frequency and luminance. *Neuropsychologia*, *38*(12), 1555–1564.

Riegel, M., Żurawski, Ł., Wierzba, M., Moslehi, A., Klocek, Ł., Horvat, M., ... & Marchewka, A. (2015). Characterization of the Nencki Affective Picture System by discrete emotional categories (NAPS BE). *Behavior research methods*, 1-13.

Rozin, P., Haidt, J., & McCauley, C. R. (2000). Disgust. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 637–653). New York, NY: Guilford Press.

Sawchuk, C. N., Lohr, J. M., Lee, T. C., & Tolin, D. F. (1999). Exposure to disgust-evoking imagery and information processing biases in blood-injection-injury phobia. *Behaviour Research and Therapy*, *37*(3), 249–257.

Schienle, A., Walter, B., Stark, R., & Vaitl, D. (2002). Ein fragebogen zur erfassung der ekelempfindlichkeit (FEE). *Zeitschrift für Klinische Psychologie und Psychotherapie; Forschung und Praxis*, *31*(2), 110–120.

Wessa, M., Kanske, P., Neumeister, P., Bode, K., Heissler, J., & Schönfelder, S. (2010). EmoPics: Subjektive und psychophysiologische Evaluationen neuen Bildmaterials für die klinisch-biopsychologische Forschung. *Zeitschrift für Klinische Psychologie und Psychotherapie, 39*(Suppl. 1/11), 77.

Wierzba, M., Riegel, M., Wypych, M., Jednoróg, K., Turnau, P., Grabowska, A., Marchewka, A. (2015). Basic Emotions in the Nencki Affective Word List (NAWL BE): New Method of Classifying Emotional Stimuli. *PLoS ONE* 10(7): e0132305. doi:10.1371/journal.pone.0132305

Wundt, W. M. (1896). *Grundriss der Psychologie*. W. Engelmann.