RESEARCH ARTICLE

Facial expressions of Asian people exposed to constructed urban forests: Accuracy validation and variation assessment

Haoming Guan[1], Hongxu Wei[2,3]*, Richard J. Hauer[4], Ping Liu[5]

1 School of Geographical Sciences, Northeast Normal University, Changchun, China, **2** Key Laboratory of Wetland Ecology and Environment, Northeast Institute of Geography and Agroecology, Chinse Academy of Sciences, Changchun, China, **3** University of Chinese Academy of Sciences, Beijing, China, **4** College of Natural Resources, University of Wisconsin-Stevens Point, Stevens Point, Wisconsin, United States of America, **5** College of Forestry, Shenyang Agricultural University, Shenyang, China

* weihongxu@iga.ac.cn

Abstract

An outcome of building sustainable urban forests is that people's well-being is improved when they are exposed to trees. Facial expressions directly represents one's inner emotions, and can be used to assess real-time perception. The emergence and change in the facial expressions of forest visitors are an implicit process. As such, the reserved character of Asians requires an instrument rating to accurately recognize expressions. In this study, a dataset was established with 2,886 randomly photographed faces from visitors at a constructed urban forest park and at a promenade during summertime in Shenyang City, Northeast China. Six experts were invited to choose 160 photos in total with 20 images representing one of eight typical expressions: angry, contempt, disgusted, happy, neutral, sad, scared, and surprised. The FireFACE ver. 3.0 software was used to test hit-ratio validation as an accuracy measurement (ac.) to match machine-recognized photos with those identified by experts. According to the Kruskal-Wallis test on the difference from averaged scores in 20 recently published papers, contempt (ac. = 0.40%, P = 0.0038) and scared (ac. = 25.23%, P = 0.0018) expressions do not pass the validation test. Both happy and sad expression scores were higher in forests than in promenades, but there were no difference in net positive response (happy minus sad) between locations. Men had a higher happy score but lower disgusted score in forests than in promenades. Men also had a higher angry score in forests. We conclude that FireFACE can be used for analyzing facial expressions in Asian people within urban forests. Women are encouraged to visit urban forests rather than promenades to elicit more positive emotions.

Introduction

The purpose of planting, growing, and managing urban tree populations is to provide the ecological services of urban forests and to promote human well-being in urban green spaces [1]. The principle of the construction of sustainable urban forests is that a community should

promote the social and ecological needs of people [2, 3]. An urban forest experience can elicit more positive emotional responses of pedestrians compared to a promenade experience [4, 5]. The efficiency of said emotional response, however, may not always be at a level that improves mental health [6, 7].

Emotion is a transitory mental response to an external stimuli [8]. Emotional response to forests can be assessed using questionnaire methodology [9]. Pilot studies in this context mainly argues that a forest experience can counter negative emotions by alleviating anxiety and stress [10–12]. However, one vital issue in these studies is that there is an absence of systematic validation on the use of questionnaires for forest visitors [13]. It is uncertain whether facial expressions were related to internal emotions and the extent to which emotions matched expression scores. Validation can solve this issue.

Emotions can be categorized as either a felt expression with an emotional cue (Duchenne) or an unfelt expression with a communicative cue (non-Duchenne) [14]. Both kinds of facial expressions can be assessed using facial expression recognition techniques [8]. For over five decades, artificial recognition of facial expressions was believed to be a reliable evaluation of emotions [15]. Automatic recognition instruments outperform traditional approaches in efficiency and accuracy [16, 17]. However, validation of facial expression recognition instruments is needed to understand the level of accuracy and precision of this recognition. It is uncertain whether facial expressions were related to internal emotions and the extent to which emotions matched expression scores. Validation can solve this issue.

The validation of face reading is a test that quantifies the percentage of agreement between artificially matching the aimed emotion [15, 18] and automatic machine reading [16, 17, 19, 20]. Early methods to validate the accuracy of recognizing emotional expressions used "matching scores" that were rated by observers regarding photos with intended expressions [15]. This was termed as "the percentage of observers who choose the predicted label" [21]. Early methods were further developed by instruments trained using datasets of expressions. Currently, notable software for facial expression recognition include Facial Action Coding System (FACS) [16, 19], FaceReader [16, 19], Affectiva Media Analytics [20], Face Analysis Computer Expression Toolbox (FACET), and iMotion [17, 20]. Datasets of facial expressions frequently used to train machine learning mostly come from the Warsaw Set of Emotional Facial Expression Pictures and the Amsterdam Dynamic Facial Expression Set [19]. The arithmetic coding for the movement of facial muscles originated from prototypical expressions on western-race faces. However, emotional expressions are not universal [22]. To more accurately detect Asian facial expressions, software that is trained by datasets of facial expressions of Asian volunteers is used [23–25]. Validation can increase the accuracy of facial expression recognition by excluding inaccurate responses. The precision of matching facial expressions to internal emotions can be increased when photos were recognized and rated by Asian volunteers.

FireFACE (Zhilunpudao Agric. S&T Inc., Changchun, China) is a software that was produced to analyze the facial expressions of Asian people [23]. The basic arithmetic was established through machine-training using 30,000 photos of facial expressions from Asian people. Version (ver.) 1.0 documented three expressions (happy, sad, and neutral) that were classified by experts [23, 25], and ver. 2.0 documented five expressions (happy, sad, neutral, angry, and scared). Ver. 3.0 was analyzed eight basic expressions (happy, sad, neutral, angry, scared, surprised, disgusted, and contempt) using the initial dataset and a subsequent dataset of photos of Asian urban forest visitors from across mainland China [24, 26]. FireFACE ver. 1.0 has been successfully used for assessing the variation of emotional expressions in university campuses [23] and at an urban forest park [25]. FireFACE ver. 3.0 has been used to detect the combined effects of geographical variation across urbanization gradient on facial expressions of urban forest visitors in Northeast China [24]. FireFACE has shown the desired precision in facial

analysis for people in urban forests, although most expressions were subtle. It is necessary to further increase matching accuracy by validating facial expression scores.

The change in setting along the urbanization gradient serves as cues for people to respond with varied emotional expressions [27, 28]. The perception to infrastructure and openness are determinants of emotional variation at different places [23, 28]. When urban forests are considered an objective infrastructure, people will show particular expressions different from those in city settings [24, 25]. In this study, FireFACE ver. 3.0 was used to test the difference in the emotional perceptions between people in constructed urban forests and people in promenades. Only facial expression that passed the validation test were used for geographical comparison. Based on the current success of subtle expressions analysis using FireFACE, we hypothesized that: (i) at least five out of the eight matching scores can meet the validation accuracy of a commercial software, and (ii) people in urban forests will show a significant difference in emotions, not only in basic expressions (happy, sad, and neutral) but also in implicit ones (angry, surprised, scared, disgusted, and contempt).

Materials and methods

Field data collection

Field data were collected from an urban forest park and an promenade in Shenyang City (41˚ 11'–42˚17' N, 122˚21'–123˚48' E). Shenyang is located in the transitional belt between Changbai Mountains and the alluvial plain of Liaohe River. Shenyang had 8.3 million permanent residents distributed across an area of 6.3 million km^2 built-up region in 2018. Shenyang is located in a semi-humid, temperate continental climate zone with annual average temperatures of 6.2–9.7 ˚C and a range of -32.9˚C and 38.4˚C. Annual rainfall in Shenyang ranged between 600 and 800 mm with a historical maximum precipitation of 716.2 mm. Yearly frost-free periods lasted for 155–180 d. Climatic data spanned from 1951 to 2018 [29].

Shenyang Expo Garden (SEG) (41˚49' N, 123˚37' E) was chosen as the site of urban forests and Shenyang Middle Street (SMS) (41˚48' N, 123˚25' E) as the promenade (Fig 1). SEG was established in February of 1959 with an openness area of 211 ha with 196 ha of green lands and 6.5 ha of watershed. Urban forests in SEG were constructed since 1988. The daily number of visitors in SEG ranged between 0.3 million and 0.7 million which is the highest record for all green spaces around the plains of Liaohe River. SMS has a length of 579.3 m and a width of 11.7 m, the longest promenade in mainland China. SMS has had a long history of use from 1625 up to now. SMS is rarely greened along the sidewalk and has areas fully occupied by groceries, markets, and plazas, which attracts anywhere from 0.4 to 2 million daily visitors. Therefore, SEG and SMS are two typical infrastructures with contrasting green spaces and constructed landscapes.

Participants

Eight students from College of Forestry, Shenyang Agricultural University were recruited as data collectors in this study. They were assembled as a group of volunteers on 19 June, 2020. All had been informed about the aim, process, and possible obstacles of the study. Only those who agreed to all the details of the study were recruited. The constant of participants has been documented as the S1 Raw data where participants provided written informed consent. Candidates with smoking or alcohol consumption habits were excluded in the recruitment. Eight students were randomly assigned to two groups with four in each group. One group investigated SEG and the other investigated SMS the first day. In the following day, places were exchanged. Two students in one group took photos and the other two asked participants for the consent of using photos for scientific work. Photographers used a camera of

Fig 1. Shenyang Expo Garden (forest) and Shenyang Middle Street (Promenade or Urban) in Shenyang City, Northeast China.

imx-586 (Sony NEC Optiarc Inc., Tokyo, Japan) with 4 million px which was embedded in the cellphone.

The Ethic Committee of the Research Group of Urban Forests and Wetlands, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, provided approval for your study. On weekends of the 20th and 21st of June, 2020, all visitors with typical Asian faces were photographed at SEG and SMS. All visitors whose faces have been photoed and recorded for this study had been informed about the aim of the study and provided their oral consents. This procedure was approved by the Ethics committee. Participants for photo collection supplied written informed consent. Visitors with faces characteristic of Chinese populations were subjects to be photographed [30]. Faces were easily identified through subjective recognition by students, but it was hard to distinguish between those of Chinese ethnicity and those of other East Asian countries. Therefore, we extended the standard to Asians in general. Both days were sunny and cloudless except for June 21st from 12:00 to 14:00. The temperature ranged between 21˚C and 32 ˚C in daytime with southwesterly winds at a velocity of Beaufort force 4 (24 km/h average speed). Photos in both sites were taken from 09:00 am to 05:00 pm (GMT+8) in accordance with the opening time of SEG. The route in SEG started at the entrance and ended at the exit with 4 repeated cycles of data collection along the sidewalks, while the route in SMS started at the northern entrance along the western side of the sidewalk in the morning and the eastern side in the afternoon to avoid building shadows.

Available photos with facial expressions

All photos that fulfilled the standard for further analysis needed to connect at least one visitor's face with the five facial organs—eyebrows, eyes, nose, mouth, and ears—no matter which angle the face was photographed from. Photos were labeled as potential candidates when only one ear can be seen, with the rest of the facial organs visible. All photos were cropped so that the subject's face is in the center and all organs are clearly exposed. A singular photo with all attributes is the best for facial expression analysis, but multiple photos with some attributes can be pieced together. A total of 2,886 photos met the criteria for further analysis. The approval for ethic statement has been documented as the S1 Raw data.

Validation of matching accuracy

A dataset of facial photos was generated from all documented photos for validation. Groups of 20 photos were selected from the pool of both SEG and SMS with each group demonstrating a particular emotion: angry, contempt, disgusted, happy, neutral, sad, scared, and surprised. A total of 160 photos were reviewed by six experts in the domain of urban ecology from four affiliations. The constant of experts for the dataset review has been documented to the S1 Raw data. The final edition of the dataset was revised according to suggestions from all experts and selections received unanimous agreement.

Validation was determined by the 'matching accuracy' variable, which is the matching percentage of the number of photos that were correctly recognized by the instrument for the predicted emotional expressions of prototypical faces for each of the 20 images [31–33]. Therefore, matching accuracy can be regarded as the percentage for correct matching. It is possible that facial photos may contain multiple expressions in different emerging values. Only the expression with highest value was considered for matching [20].

Give that the matching accuracy of validation for facial expressions varied widely depending on the choice of database, methodology, and instrument, we established a set of standards to screen for the validation of each of the eight expressions in our photos from SEG or SMS. A combination of the keywords of 'validation' + 'accuracy' + 'facial expression' was checked in the search engine of Web of Science (Clarivate Analytics, Philadelphia, Pennsylvania, USA). The 20 most relevant studies with specific sources of data (either from figure, words, or tables) were documented for data extraction. The criteria when screening for usage was adapted as the mean of 20 studies (Table 1) [16, 30–48]. The specific process is shown in Fig 2. Only expressions that passed the criteria were used for further assessment.

Assessment of variation and statistics

The eight students were invited again to create demographic categories. All photos were classified by gender (man vs woman) and age (senior [over 60 years-old], middle-aged [35–50 years-old], youth [15–25 years-old], toddler [0–5 years-old]). We categorized gender according to visually identifiable, biological characteristics. Age was categorized by empirical identification to visual standards of the median of each age group. Some age categorizations were ambiguous for identification between senior and middle ages or between young and toddler ages. All eight students assembled together, discussed, and voted for the final choice.

Each of the validated expressions, the dependent variable, was analyzed in response to the combined independent variables of gender ($n = 2$), age ($n = 4$), and location ($n = 2$). Only facial expressions that passed the validation with no difference from the average in other literature was used for next-step analysis.

Data were analyzed using SAS software (STS Institute, Cary, NC, USA). A new parameter, termed positive response index (PRI) [23–25], was employed to evaluate the net difference

Table 1. Summary of matching accuracy for validation studies on facial expressions.

Source	Instrument	Data source	Matching accuracy (%)							
			Ang.[1]	Con.[2]	Dis.[3]	Hap.[4]	Neu.[5]	Sad	Sca.[6]	Sur.[7]
Banziger et al. [34]	Artificial recognition	Figure	27	-	85	56	-	47	-	-
Ebner et al. [36]	Artificial recognition	Figure	80	-	69	-	84	73	84	-
Besel and Yuille [35]	Artificial recognition	Table	82	-	55	69	75	58	54	51
Matuszewski et al. [37]	Eigenface, Fisherface	Table	62	-	55	83	-	51	35	62
Huang et al. [30]	E-Prime	Figure	54	-	32	63	43	33	27	21
Maniglio et al. [38]	Artificial recognition	Table	50	-	-	93	65	75	65	-
Zhang et al. [40]	Artificial recognition	Table	43	-	-	69	-	43	-	-
Olszanowski et al. [39]	Artificial recognition	Figure	95	-	100	78	57	95	78	76
Wingenbach et al. [41]	E-Prime	Figure	74	35	65	84	89	78	62	92
Vaiman et al. [43]	FACS	Table	85	-	80	95	87	78	63	91
Kim et al. [42]	FACS	Figure	87	-	63	97	93	84	50	93
Saeed et al. [46]	AERS [8], GSNMF [9], TPTSR [10]	Figure	60	-	52	54	-	57	52	50
Mishra et al. [44]	E-Prime	Figure	80	-	79	96	88	86	80	88
Prada et al. [45]	Artificial recognition	Table	79	-	69	89	75	70	58	80
Chung et al. [47]	Artificial recognition	Figure	90	-	60	99	73	83	36	92
Verpaalen et al. [48]	Artificial recognition	Figure	59	34	53	93	63	70	47	66
Skiendziel et al. [16]	FaceReader Ver. 7	Words	76	-	92	96	94	86	82	94
Yang et al. [33]	Amazon Rekognition, Baidu Research, Face++, Microsoft Azure, Affectiva	Table	64	-	91	99	98	85	56	97
Krumhuber et al. [32]	FACET	Figure	48	-	57	88	-	54	34	46
Bijsterbosch et al. [31]	Artificial recognition	Table	77	39	68	92	81	75	63	64
Mean			68.60	36.00	68.06	83.53	77.67	69.05	57.00	72.69
Standard deviation			17.41	2.16	16.70	14.05	14.90	16.51	16.71	21.57

Note:

[1] Ang., angry;

[2] Con., contempt;

[3] Dis., disgusted;

[4] Hap., happy;

[5] Neu., neutral;

[6] Sca., scared;

[7] Sur., surprised;

[8] AERS, automatic expression recognition system;

[9] GSNMF, graph-preserving sparse nonnegative matrix factorization;

[10] TPTSR, two-phase test sample representation.

between happy and sad expression scores. In validation, the Kruskal-Wallis test was repeatedly used to detect the difference between the critical standard for expressions from literature ($n = 20$) and expressions from our database. The basic probability of significance was taken at the 0.05 level as adjusted by the Bonferroni method to 0.00625 due to eight repeated comparisons. Thus, scores for contempt, neutral, scared, and surprised expressions were not recorded in the 20 documented studies for validation (Table 1). Only the expressions that did not show significant difference between our database and previous studies were used as parameters for further assessment. In variation assessment, all data were ranked to avoid abnormally
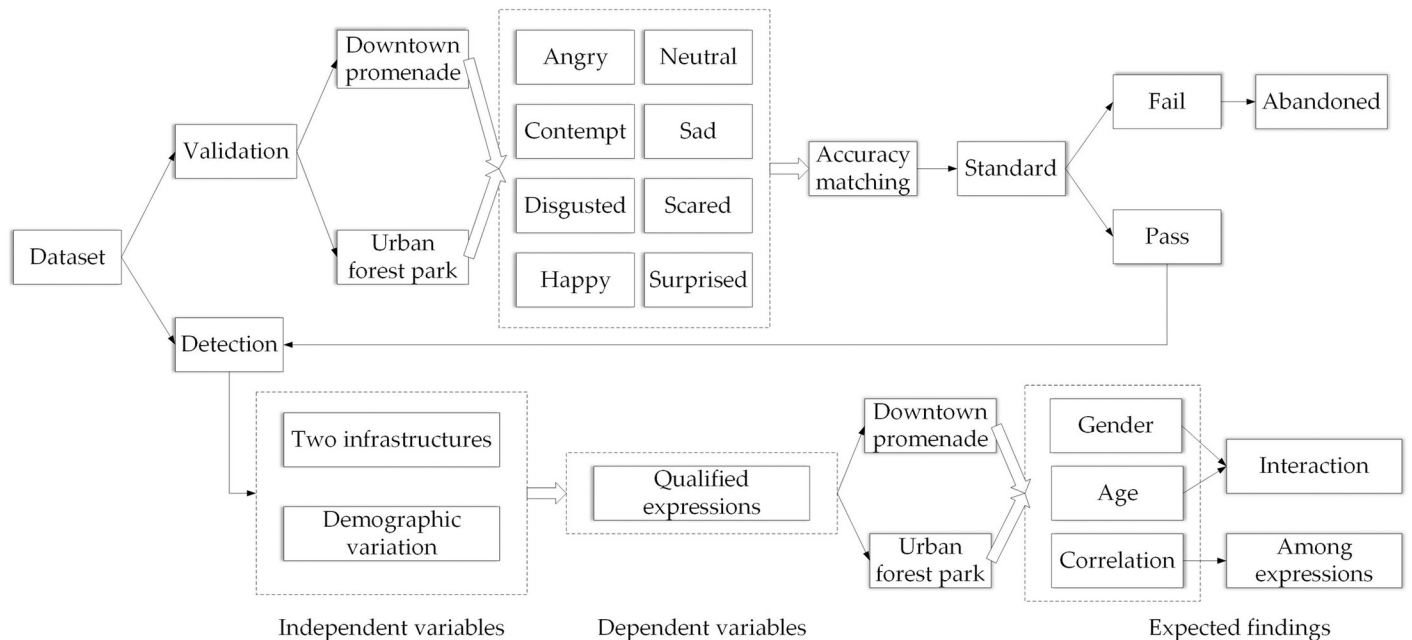
Fig 2. The whole process of the layout of whole study from validation to analysis.

distributed data that invalidated use of a general linear model. Every expression was tested for response to three-way analysis of variance (ANOVA) across gender, age, and location. When significant effect was found, data were ranked and compared by a one-way ANOVA with all combined factors together as the single source of variance ($\alpha = 0.05$). The principle component analysis (PCA) was used to bridge the grouped tendency of correlation that had been used in several former studies [49, 50].

Results

Validation of recognition accuracy

As shown in Table 1, the selected 20 publications reporting facial expression accuracy did not supply data for all eight expressions. For example, only three out of the 20 publications evaluated accuracy for contempt expression and 15 out of 20 for neutral expression. The highest accuracy was found for happy expression scores, followed by neutral and surprised expressions. The lowest accuracy was found for contempt expression scores, and the rest were all above 50%.

Results on FireFACE's accuracy in recognizing the eight facial expressions are shown in Table 2. The accuracy in recognizing contempt and scared expressions was significantly lower than the average from the 20 publications. Although averaged accuracy, when comparing our database to historical ones, was lower by 18–69% for the rest of the facial expressions, repeated Kruskal-Wallis tests did not indicate any significant difference because raw data showed a large variation in both our and previous databases. Therefore, we accept FireFACE's accuracy in recognizing facial expressions, but only for the six expressions of anger, disgust, happiness, sadness, neutral, and surprise.

Analysis of variance on facial expressions in Shenyang

As shown in Table 3, excluding scores for anger and PRI, scores for the rest of the facial expressions demonstrated significant responses to variation between forest and urban locations. All

Table 2. Accuracy of eight facial expressions using FireFACE software and repeated Kruskal-Wallis tests on the difference from records in documented studies that are shown in Table 1.

Facial expressions	DF	Accuracy (%)	SD [1]	Chi-Square	Pr > Chi-Square [2]
Angry	1	56.11	13.81	5.5388	0.0186
Contempt	1	0.40	0.90	8.3623	0.0038*
Disgusted	1	43.98	22.72	6.0610	0.0138
Happy	1	60.55	35.24	1.0296	0.3102
Neutral	1	23.85	20.89	7.3023	0.0069
Sad	1	55.53	11.69	7.1742	0.0074
Scared	1	25.23	20.57	9.6959	0.0018*
Surprised	1	47.57	15.75	2.8156	0.0934

Note:

[1] SD, standard deviation;

[2] asterisk indicates significant difference at the 0.00625 level.

https://doi.org/10.1371/journal.pone.0253141.t002

facial expression scores, except scores for surprise, were different by gender. Happy expression scores did not vary by visitor age, but responded significantly to the interaction between location and age. Happy expression scores also responded to the interaction between gender and age. In addition, angry, surprised, and disgusted expression scores showed a significant response to the interaction between location and gender as well.

Response of happy expression scores

Women had higher happy expression scores than men in both urban forest and promenade (Fig 3A). Women had the highest happy expression score in the forest locations. Youths and senior women had higher happy expression scores than toddlers and senior men (Fig 3B).

Forest visitors had higher happy expression scores than those in promenades by 13%. Women had higher happy expression scores than men by 8%. Young visitors had higher happy expression scores than toddlers by 12%. The happy expression scores of youths were not different from middle-aged or old-aged visitors.

Responses of angry, surprised, and disgusted expression scores

Men in forest areas had the highest angry expression score in response to interaction between gender and location (Fig 4A). In contrast, women in promenades had the highest surprised

Table 3. P values from analysis of variance (ANOVA) of location (L), gender (G), age (A), and their interactions on ranked scores of neutral, happy, sad, angry, surprised, and disgusted expressions.

Source	DF [1]	Neutral	Happy	Sad	PRI [2]	Angry	Surprised	Disgusted
L	1	< .0001 [3]	< .0001	0.0002	0.5283	0.0978	< .0001	0.0005
G	1	0.0046	0.0022	< .0001	0.0002	0.0455	0.6032	< .0001
A	9	< .0001	0.0704	< .0001	< .0001	0.0001	< .0001	< .0001
L×G	1	0.2388	0.8089	0.1089	0.1537	0.0082	0.0406	0.0151
L×A	3	0.7786	0.0083	0.0771	0.2625	0.5605	0.5979	0.4644
G×A	3	0.1385	0.0467	0.6569	0.9874	0.6179	0.5579	0.9863
L×G×A	3	0.8708	0.5846	0.2112	0.1217	0.1611	0.2955	0.2107

Note:

[1] DF, degree of freedom;

[2] PRI, positive response index;

[3] Bold font highlights significant effect.

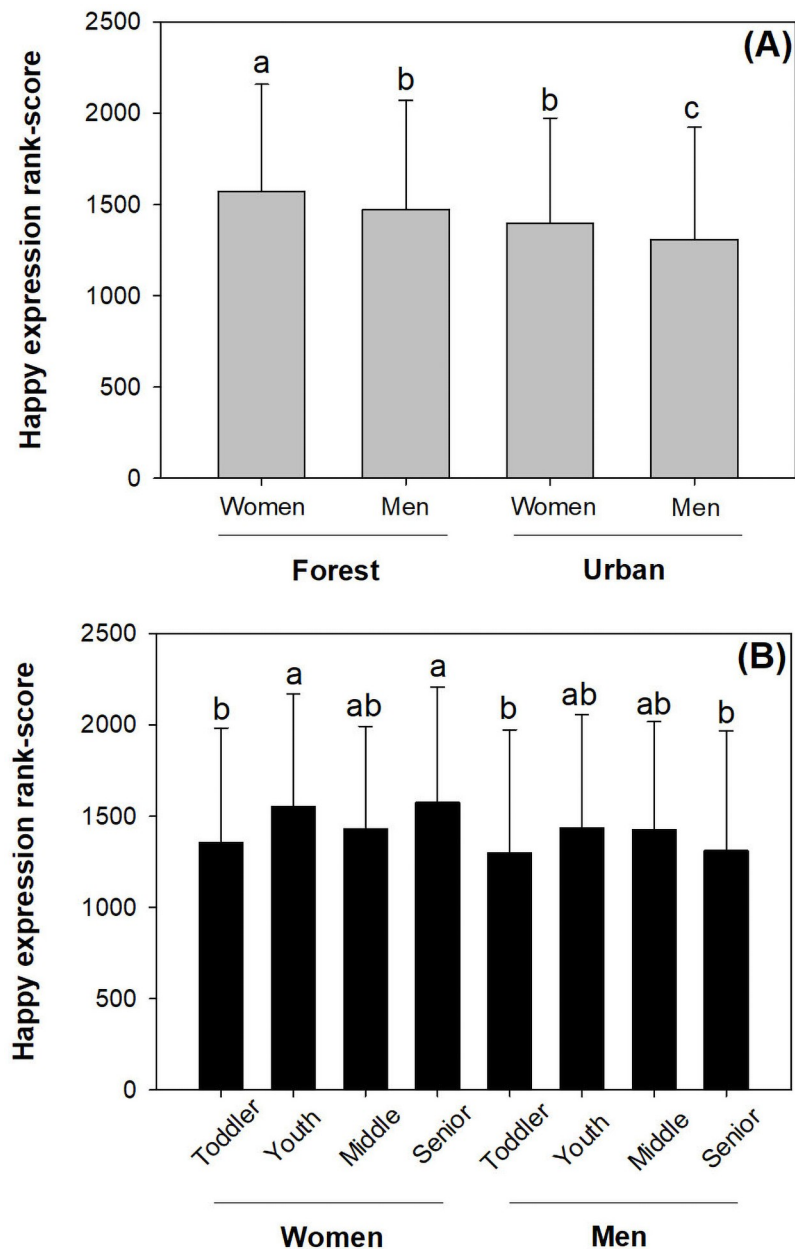https://doi.org/10.1371/journal.pone.0253141.t003

Fig 3. Happy expression scores for men and women at different ages in forest and urban locations. Scores have been transformed by ranking. Bars stand for standard errors. Different letters indicate significant difference according to Duncan test at 0.05 level.

and disgusted expression scores (Fig 4B and 4C). Although the surprised expression scores of men in promenades were lower than women, men's scores were higher than those of women in forest areas (Fig 4B).

There were no distinct angry expression scores between visitor in different locations (Table 3). Women generally had a higher angry expression score than men by 4%. Middle-aged visitors had higher angry scores than toddlers and senior citizens by 22% and 11%, respectively.
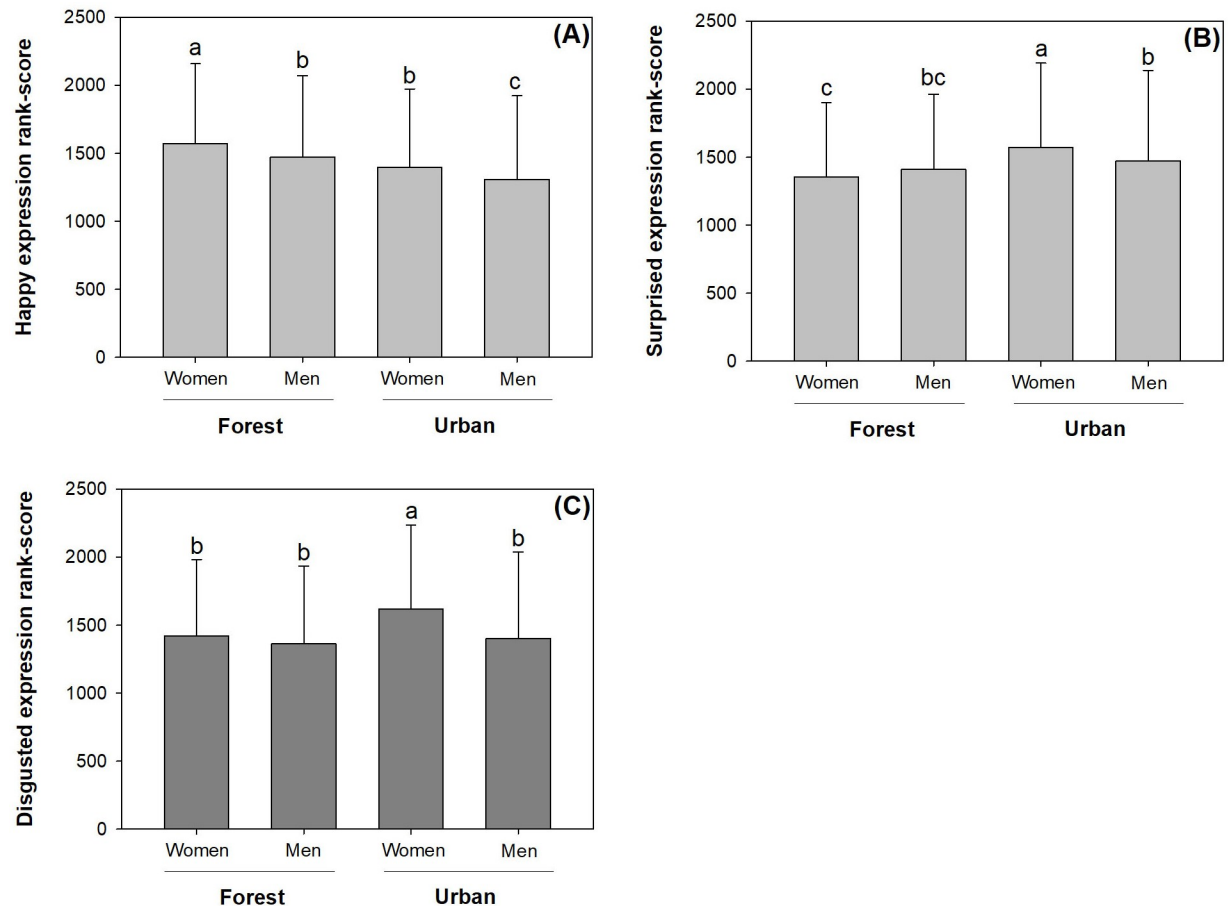
Fig 4. Expression scores for angry, surprised, and disgusted emotions in men and women in forest and urban locations. Scores have been transformed by ranking. Bars stand for standard errors. Different letters indicate significant difference according to Duncan test at 0.05 level.

Both surprised and disgusted expression scores were higher in promenade than in forests. Both expression scores were also higher for youths than for toddlers and senior citizens. Women had higher disgusted expression scores than men by 9%.

Responses of neutral and sad expression scores

Both neutral and sad expression scores were higher in the urban forests than in cities (Fig 5A and 5B). Women had higher neutral and sad expression scores than men (Fig 5C and 5D). The neutral expression scores generally decreased as age increased, but older visitors had higher sad expression scores than toddlers and young visitors (Fig 5E and 5F).

Response of PRI

PRI did not show any significant response to the difference between forest and urban locations (Table 3). Men had higher PRI than women by 8%. Toddlers and youths had higher PRI than older visitors.

PCA analysis

The data pool of the six facial expressions and PRI combined indicate a variation in which the first PC explains 35.3% and the first two PC cumulatively explains 53.93%. Therefore, an
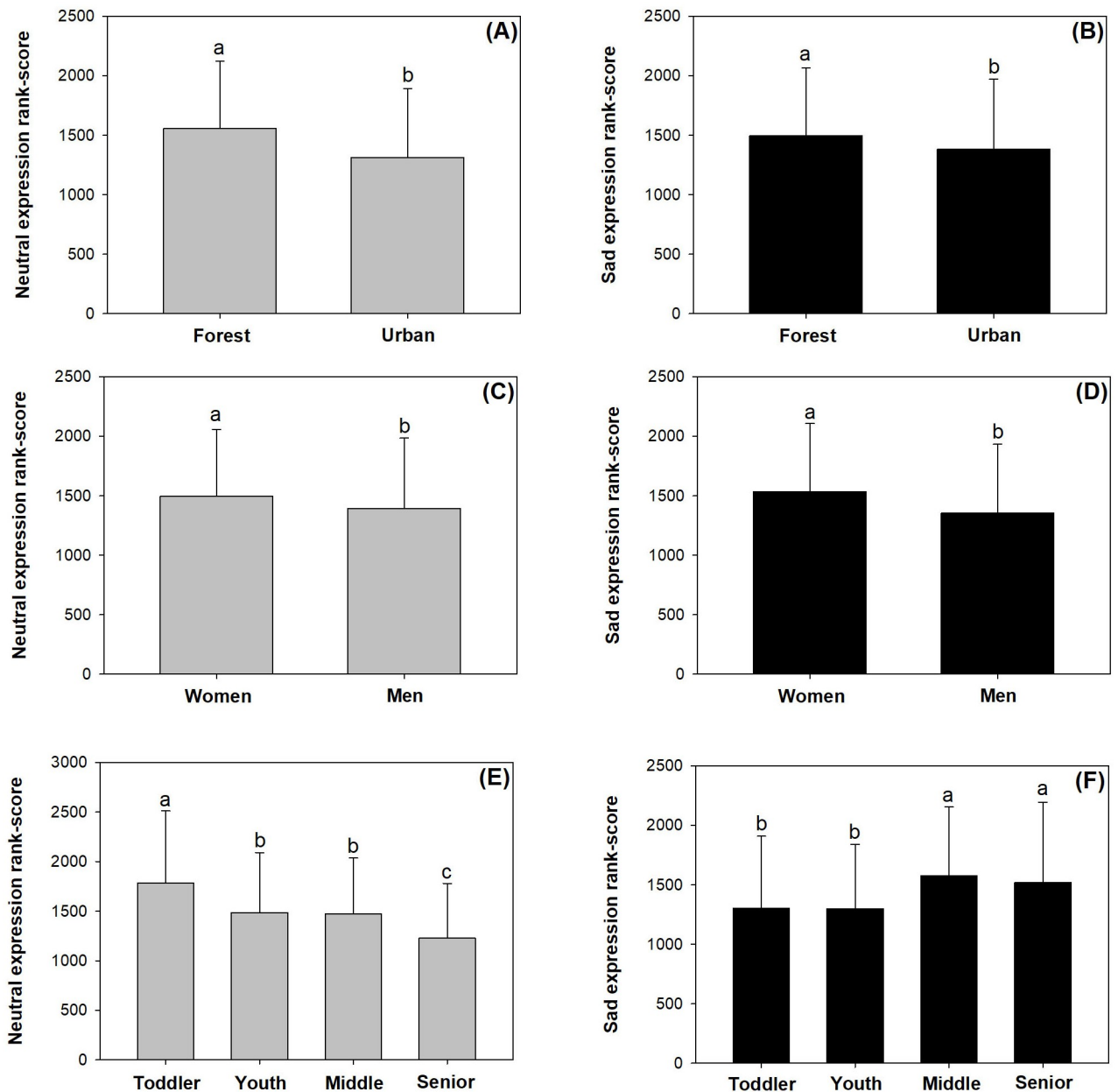
Fig 5. Neutral and sad scores with variations by location (forest vs urban), gender (men vs women), and age (toddler, youth, middle, senior). Scores have been transformed by ranking. Bars stand for standard errors. Different letters indicate significant difference according to Duncan test at 0.05 level.

explanation above 50% supports a further analysis on the synthesis of the first two PCs. The sad expression score had an inverse relationship with the happy expression score and PRI (Fig 6). Both the disgusted and angry expression scores had an inverse relationship with the neutral expression score, but the relationship was weaker for anger than for disgust. The surprised expression score did not show any obvious relationship with any other emotional expressions.
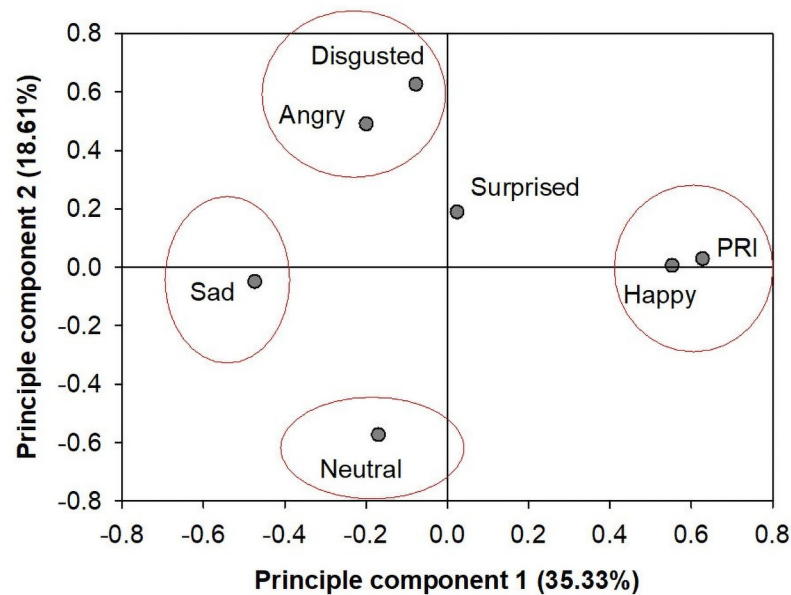
Fig 6. Eigenvalues of happy, sad, neutral, angry, disgusted, and surprised expression scores and positive response index (PRI) in the first two principle components.

https://doi.org/10.1371/journal.pone.0253141.g006

Discussion

Validation of recognizing accuracy

In our study, contempt and scared expression scores evaluated by FireFACE failed to pass the validation test. The scared expression was also recognized at a low accuracy for the faces of Chinese people even when using the three-dimensional paradigm technique [30]. The contempt emotion is difficult to detect, not only with FireFACE but also with other instruments with a wider range of users. This is further corroborated by the fact that only three out of the recent 20 relevant publications had accuracy in recognizing the contempt expression. The matching accuracy in these three cases were around 30%, which was much lower than the recognition of other expressions. In addition, two cases of accuracies were achieved through artificial rating [31, 48], and only one case published a matching accuracy that was given by an instrument, E-Prime [41]. Therefore, further improvement is needed for FireFACE to recognize the contempt expression because scores were too low for accurate determination. Machine learning techniques need improvement to recognize the exhibition of the contempt emotion in different groups of people.

Chinese people, characteristic of how Asians typically display emotions, show emotions implicitly rather than explicitly. For example, Chinese people's expression of fear forms more slowly than other negative expressions when depicting pain, even after sad emotions priming [51]. Chinese people's tendency to suppress expressions of fear by adopting a self-reserved character can be extended to the Korean population's Yonsei database [47]. Krumhuber et al. [32] compared human and machine (FACET software) validations across 14 datasets of dynamic facial expressions and only obtained a 34% accuracy in recognizing a scared expression. They further found that the scared expression was easily confused with the surprised expression. Matuszewski et al. [37] checked a dataset of facial expressions from 80 clinic patients and, again, found a low recognition of the scared expression and corroborated the easy confusion with the surprised expression. Matuszewski et al. [37] further compared

different levels of the scared scores and indicated that patients directly expressed fear only when exposed to extreme pain. Otherwise, they would choose to reserve their expressions to avoid perception by others. Overall, more precisely distinguishing between scared and surprise expressions is suggested to increase the accuracy of recognition.

Our matching accuracy was generally lower than those found in previous studies. This can be explained with two reasons. First, our dataset that was used to train FireFACE contained subjective errors when artificially documenting different facial photos into any of the eight types of expressions. Second, our objects receiving the tests were collected by randomly photographing visitors and subjectively labeling the type of facial expressions, hence, the precision was limited in addition to subjective error. In contrast, both machine-training and objects-testing in the 20 studies reviewed in this paper employed models with instructions to exhibit aimed expressions. Even so, our results on the matching accuracies for the six facial expressions of anger, disgust, happiness, sadness, neutral, and surprise were not statistically different from current ones. Therefore, FireFACE recognition for Asian facial expressions has an acceptable accuracy for six expressions. Thus, we can accept our first hypothesis.

Facial expression in constructed forests and promenade

It is unexpected that both happy and sad expressions were higher for visitors in constructed forests than in urban promenades. However, there was no difference in net positive score between the two locations that differed in greenspace. Therefore, the experience in the forest did not result in expressions that were extremely different from experience in the promenade. Instead, forest visitors showed less disgusted expressions than those in the promenade. Disgust is a type of negative emotion that is less extreme than sadness. Therefore, our results confer findings that more negative emotions can be reduced in forests than in built-up regions [10–12] according to the reason that people in cities elicited more negative emotions in the form of disgust than compared to those in forests. Wei et al. [24] also found that people in an urban forest park near the center of a city would show more negative expressions than those in forest parks in remote rural regions. Eigenvalues of the neutral scores were positively correlated with happy scores and both were inverse to sad scores, which suggests that people generally perceived net positive emotion scores in forests. Results of Wei et al. [25] concur with our finding in that people in forests showed more happiness than in an urban street and PRI. We can accept our second hypothesis.

Both men and women showed less happy expressions in urban locations than in forests, supporting the above-mentioned higher happy scores in the forest. Women also showed more positive emotions from a forest experience than in the city. These results concur with findings of Wei et al. [24] who also reported that women showed more positive expressions than men. Our results also revealed that women in forests showed lower surprised and disgusted expressions than those in the city. Negative expressions of men were not statistically different. These findings suggest that women are more sensitive and responsive than men, and thus showed more positive and negative expressions in forests and promenades, respectively. This corroborates a previous investigation which indicated that healthy men showed reduced emotion processing efficiency relative to women [52]. However, men reacted to forest locations by processing highly efficient angry expressions while women did show any differences in angry expressions between urban and forest locations. This did not lead to a divergence of angry expressions between urban and forest locations; hence men's higher angry score in forests was the result of their psychological response. A review using a functional-evolutionary analysis indicated that it is more advantageous for men to show angry facial expression as it signals dominance, averts aggression, and deters mate poaching; it is more advantageous for women

to display happy facial expressions as it signals their willingness for childcare, tending, and befriending [53].

There was a negative relationship between neutral expressions and disgusted and surprised expressions. There was a higher neutral expressions score in forests than in promenades in accordance with less disgust and surprise expressions. From this, we concluded that people in forests are more calm than in the city, whereas the excited emotion was easily confused with, and represented as, a disgusted expression. It is the peaceful environment with abundant green color and moisture in the forests that caused the calm feeling expressed as neutral faces for visitors in forests [25].

Limits of the study

Because of accuracy, we excluded ratings on contempt and scared expressions from the analysis. It may be hard to obtain desired accuracy in recognizing contempt expressions until deep learning techniques can improve upon current limitations. A scared emotion is an important facial expression that frequently emerges in daily life. Some other instruments, such as E-Prime [44] and FaceReader [16], were reported to recognize fear in face reading at an accuracy as high as 80%, and deserves to be used in future studies to test the hit ratio for our database on Asian faces in urban forest parks. We turned to experts to help classify the typical facial expressions from our dataset. This is a useful approach to document different types of expressions. However, a more reasonable way is to classify the facial expressions by running the same photo through different software. The average score across results from different machines will provide a more reliable identification of emotions than human perception can. Furthermore, we did not discuss the interaction between gender and age on facial expressions because they do not match the theme of this study and our study results cannot support a deeper analysis on this relationship. Finally, only two locations were employed in this study, which is enough to support a frontier study on validation and assessment. Although we employed practical methodology with validation, our data still may suffer uncertainties from collection and unexpected errors. It is likely that more tests on dataset from more cities and locations would increase accuracy with the results. The inaccuracy of matching human and machine-recognized scores will be reduced with the increase of object number. Future work is encouraged to build upon this study and include additional demographic data.

Conclusions

We compared the accuracies to match the recognized facial expressions given by FireFACE with those assessed by other instruments or artificial approach. Facial expressions of angry, disgusted, happy, sad, neutral, and surprised emotions passed the validation test because their scores were within a level of statistical acceptance. However, contempt and scared expression scores were too low and these were excluded from further analysis. We collected a total of 2,886 photos from visitors in constructed urban forests and in a promenade during summertime in Shenyang, Northeast China. There were no extreme differences in emotional expressions between forest and urban locations. A gender interaction with location showed that women exhibited more positive, but less negative, expressions in forests than in the promenade. Synthesizing these findings, we suggest that women visit constructed urban forest parks more often to elicit greater happiness and decreased negative emotions.

Supporting information

S1 Raw data.
(PDF)

Acknowledgments

Authors acknowledge Prof. Yutao Wang from Shenyang Agricultural University, Prof. Shenglei Guo from Heilongjiang University of Chinese Medicine, and Dr. Qiao Li from Changchun Institute of Technology for their participation to review and revise the dataset for the validation of matching accuracy in addition to the effort from those in author board.

Author statement

The validation part of this study has been published as preprint which can be found in the following link: https://www.preprints.org/manuscript/202010.0265/v1.

Author Contributions

References

1. Foo CH. Linking forest naturalness and human wellbeing-A study on public's experiential connection to remnant forests within a highly urbanized region in Malaysia. Urban For Urban Green. 2016; 16: 13–24.

2. Yang J, Luo X, Jin C, Xiao X, Xia J. Spatiotemporal patterns of vegetation phenology along the urban–rural gradient in Coastal Dalian, China. Urban For Urban Green. 2020; 54: 126784.

3. Yang J, Sun J, Ge Q, Li X. Assessing the impacts of urbanization-associated green space on urban land surface temperature: A case study of Dalian, China. Urban For Urban Green. 2017; 22: 1–10.

4. Sonti NF. Ambivalence in the Woods: Baltimore resident perceptions of local forest patches. Soc Natur Resour. 2020; 33: 823–841.

5. Joung D, Kim G, Choi Y, Lim H, Park S, Woo JM, et al. The Prefrontal cortex activity and psychological effects of viewing forest landscapes in autumn season. Int J Env Res Pub Health. 2015; 12: 7235–7243. https://doi.org/10.3390/ijerph120707235 PMID: 26132477

6. Akpinar A, Barbosa-Leiker C, Brooks KR. Does green space matter? Exploring relationships between green space type and health indicators. Urban For Urban Green. 2016; 20: 407–418.

7. Wolf KL, Lam ST, McKeen JK, Richardson GRA, van den Bosch M, Bardekjian AC. Urban trees and human health: A scoping review. Int J Env Res Pub Health. 2020; 17: 30. https://doi.org/10.3390/ijerph17124371 PMID: 32570770

8. Ekman P. Facial expression and emotion. Am Psychol. 1993; 48: 384–392. PMID: 8512154

9. Takayama N, Saito H, Fujiwara A, Horiuchi M. The effect of slight thinning of managed coniferous forest on landscape appreciation and psychological restoration. Prog Earth Planet Sci. 2017; 4: 15.

10. Bielinis E, Takayama N, Boiko S, Omelan A, Bielinis L. The effect of winter forest bathing on psychological relaxation of young Polish adults. Urban For Urban Green. 2018; 29: 276–283.

11. Badiu DL, Ioja CI, Patroescu M, Breuste J, Artmann M, Nita MR, et al. Is urban green space per capita a valuable target to achieve cities' sustainability goals? Romania as a case study. Ecol Indic. 2016; 70: 53–66.

12. Hauru K, Lehvavirta S, Korpela K, Kotze DJ. Closure of view to the urban matrix has positive effects on perceived restorativeness in urban forests in Helsinki, Finland. Landscape Urban Plan. 2012; 107: 361–369.

13. Aerts R, Honnay O, Van Nieuwenhuyse A. Biodiversity and human health: Mechanisms and evidence of the positive health effects of diversity in nature and green spaces. Brit Med Bull. 2018; 127: 5–22. https://doi.org/10.1093/bmb/ldy021 PMID: 30007287

14. Surakka V, Hietanen JK. Facial and emotional reactions to Duchenne and non-Duchenne smiles. Int J Psychophysiol. 1998; 29: 23–33. https://doi.org/10.1016/s0167-8760(97)00088-3 PMID: 9641245

15. Ekman P, Sorenson ER, Friesen WV. Pan-cultural elements in facial displays of emotion. Science. 1969; 164: 86–88. https://doi.org/10.1126/science.164.3875.86 PMID: 5773719

16. Skiendziel T, Rosch AG, Schultheiss OC. Assessing the convergent validity between the automated emotion recognition software Noldus FaceReader 7 and Facial Action Coding System Scoring. PLoS One. 2019; 14: 18. https://doi.org/10.1371/journal.pone.0223905 PMID: 31622426

17. Calvo MG, Fernandez-Martina A, Recio G, Lundqvist D. Human observers and automated assessment of dynamic emotional facial expressions: KDEF-dyn database validation. Front Psychol. 2018; 9: 12.

18. Langner O, Dotsch R, Bijlstra G, Wigboldus DHJ, Hawk ST, van Knippenberg A. Presentation and validation of the Radboud Faces Database. Cogn Emot. 2010; 24: 1377–1388.

19. Lewinski P, den Uyl TM, Butler C. Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. J Neurosci Psychol E. 2014; 7: 227–236.

20. Stockli S, Schulte-Mecklenbeck M, Borer S, Samson AC. Facial expression analysis with AFFDEX and FACET: A validation study. Behav Res Methods. 2018; 50: 1446–1460. https://doi.org/10.3758/s13428-017-0996-1 PMID: 29218587

21. Nelson NL, Russell JA. Universality revisited. Emotion Review. 2013; 5: 8–15.

22. Jack RE, Garrod OGB, Yu H, Caldara R, Schyns PG. Facial expressions of emotion are not culturally universal. PNAS. 2012; 109: 7241–7244. https://doi.org/10.1073/pnas.1200155109 PMID: 22509011

23. Wei H, Hauer RJ, Zhai X. The relationship between the facial expression of people in university campus and host-city variables. Appl Sci. 2020; 10: 17.

24. Wei H, Hauer RJ, Chen X, He X. Facial expressions of visitors in forests along the urbanization gradient: What can we learn from selfies on social networking services? Forests. 2019; 10: 14.

25. Wei H, Ma B, Hauer RJ, Liu C, Chen X, He X. Relationship between environmental factors and facial expressions of visitors during the urban forest experience. Urban For Urban Green. 2020; 53: 10.

26. Wei H, Hauer RJ, He X. A forest experience does not always evoke positive emotion: A pilot study on unconscious facial expressions using the face reading technology. Forest Policy and Economics. 2021; 123: 102365.

27. Ali G, Ali A, Ali F, Draz U, Majeed F, Yasin S, et al. Artificial Neural Network Based Ensemble Approach for Multicultural Facial Expressions Analysis. Ieee Access. 2020; 8: 134950–134963.

28. Kang YH, Jia QY, Gao S, Zeng XH, Wang YY, Angsuesser S, et al. Extracting human emotions at different places based on facial expressions and spatial clustering analysis. T Gis. 2019; 23: 450–480.

29. Baidu Cyclopedia. 2020. Shenyang City (The capital of Liaoning province, deputy provincial city) 2020. cited [11 June 2020]. In: Baidu Cyclopedia. https://baike.baidu.com/item/%E6%B2%88%E9%98%B3/13034?fromtitle=%E6%B2%88%E9%98%B3%E5%B8%82&fromid=124784&fr=aladdin

30. Huang CLC, Hsiao S, Hwu HG, Howng SL. The Chinese facial emotion recognition database (CFERD): A computer-generated 3-D paradigm to measure the recognition of facial emotional expressions at different intensities. Psych Res. 2012; 200: 928–932. https://doi.org/10.1016/j.psychres.2012.03.038 PMID: 22503384

31. Bijsterbosch G, Mobach L, Verpaalen IAM, Bijlstra G, Hudson JL, Rinck M, et al. Validation of the child models of the Radboud Faces Database by children. Int J Behav Dev. 2020: 7.

32. Krumhuber EG, Kuster D, Namba S, Skora L. Human and machine validation of 14 databases of dynamic facial expressions. Behav Res Methods. 2020: 16.

33. Yang KN, Wang CF, Sarsenbayeva Z, Tag B, Dingler T, Wadley G, et al. Benchmarking commercial emotion detection systems using realistic distortions of facial image datasets. Visual Comp. 2020: 20.

34. Banziger T, Grandjean D, Scherer KR. Emotion recognition from expressions in face, voice, and body: The multimodal emotion recognition test (MERT). Emotion. 2009; 9: 691–704. PMID: 19803591

35. Besel LDS, Yuille JC. Individual differences in empathy: The role of facial expression recognition. Pers Indiv Indiv Differ. 2010; 49: 107–112.

36. Ebner NC, Riediger M, Lindenberger U. FACES-A database of facial expressions in young, middle-aged, and older women and men: Development and validation. Behav Res Methods. 2010; 42: 351–362. https://doi.org/10.3758/BRM.42.1.351 PMID: 20160315

37. Matuszewski BJ, Quan W, Shark LK, McLoughlin AS, Lightbody CE, Emsley HCA, et al. Hi4D-ADSIP 3-D dynamic facial articulation database. Image Vision Comput. 2012; 30: 713–727.

38. Maniglio R, Gusciglio F, Lofrese V, Murri MB, Tamburello A, Innamorati M. Biased processing of neutral facial expressions is associated with depressive symptoms and suicide ideation in individuals at risk for major depression due to affective temperaments. Compr Psychiat. 2014; 55: 518–525. https://doi.org/ 10.1016/j.comppsych.2013.10.008 PMID: 24238931

39. Olszanowski M, Pochwatko G, Kuklinski K, Scibor-Rylski M, Lewinski P, Ohme RK. Warsaw set of emotional facial expression pictures: a validation study of facial display photographs. Front Psychol. 2015; 5: 8. https://doi.org/10.3389/fpsyg.2014.01516 PMID: 25601846

40. Zhang F, Parmley M, Wan XA, Cavanagh S. Cultural differences in recognition of subdued facial expressions of emotions. Motiv Emotion. 2015; 39: 309–319.

41. Wingenbach TSH, Ashwin C, Brosnan M. Validation of the Amsterdam dynamic facial expression set—bath intensity variations (ADFES-BIV): A set of videos expressing low, intermediate, and high intensity emotions. PLoS One. 2016; 11: 28.

42. Kim SM, Kwon YJ, Jung SY, Kim MJ, Cho YS, Kim HT, et al. Development of the Korean facial emotion stimuli: Korea university facial expression collection 2nd edition. Front Psychol. 2017; 8: 11.

43. Vaiman M, Wagner MA, Caicedo E, Pereno GL. Development and validation of an Argentine set of facial expressions of emotion. Cogn Emot. 2017; 31: 249–260. https://doi.org/10.1080/02699931. 2015.1098590 PMID: 26479048

44. Mishra MV, Ray SB, Srinivasan N. Cross-cultural emotion recognition and evaluation of Radboud faces database with an Indian sample. PLoS One. 2018; 13: 19. https://doi.org/10.1371/journal.pone. 0203959 PMID: 30273355

45. Prada M, Garrido MV, Camilo C, Rodrigues DL. Subjective ratings and emotional recognition of children's facial expressions from the CAFE set. PLoS One. 2018; 13: 21. https://doi.org/10.1371/journal. pone.0209644 PMID: 30589868

46. Saeed S, Mahmood MK, Khan YD. An exposition of facial expression recognition techniques. Neural Comput Appl. 2018; 29: 425–443.

47. Chung KM, Kim S, Jung WH, Kim Y. Development and validation of the Yonsei face database (YFace DB). Front Psychol. 2019; 10: 18.

48. Verpaalen IAM, Bijsterbosch G, Mobach L, Bijlstra G, Rinck M, Klein AM. Validating the Radboud faces database from a child's perspective. Cogn Emot. 2019; 33: 1531–1547. https://doi.org/10.1080/ 02699931.2019.1577220 PMID: 30744534

49. Wang R, Wang Y, Su Y, Tan JH, Luo XT, Li JY, et al. Spectral Effect on Growth, Dry Mass, Physiology and Nutrition in Bletilla striata Seedlings: Individual Changes and Collaborated Response. Int J Agric Biol. 2020; 24: 125–132.

50. Zhou CW, Yan LB, Yu LF, Wei HX, Guan HM, Shang CF, et al. Effect of Short-term Forest Bathing in Urban Parks on Perceived Anxiety of Young-adults: A Pilot Study in Guiyang, Southwest China. Chin Geogr Sci. 2019; 29: 139–150.

51. Song J, Wei YQ, Ke H. The effect of emotional information from eyes on empathy for pain: A subliminal ERP study. PLoS One. 2019; 14: 15. https://doi.org/10.1371/journal.pone.0226211 PMID: 31834900

52. Weisenbach SL, Rapport LJ, Briceno EM, Haase BD, Vederman AC, Bieliauskas LA, et al. Reduced emotion processing efficiency in healthy males relative to females. Soc Cogn Affect Neurosci. 2014; 9: 316–325. https://doi.org/10.1093/scan/nss137 PMID: 23196633

53. Tay PKC. The adaptive value associated with expressing and perceiving angry-male and happy-female faces. Front Psychol. 2015; 6: 6.