

RESEARCH ARTICLE

REVISED Affective rating of audio and video clips using the

EmojiGrid [version 2; peer review: 2 approved]

Alexander Toet^{1,2}, Jan B. F. van Erp^{1,3}

¹Perceptual and Cognitive Systems, TNO, Soesterberg, 3769DE, The Netherlands

²Department of Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, 3584 CS, The Netherlands ³Research Group Human Media Interaction, University of Twente, Enschede, 7522 NH, The Netherlands

V2 First published: 11 Aug 2020, 9:970 https://doi.org/10.12688/f1000research.25088.1 Latest published: 06 Apr 2021, 9:970 https://doi.org/10.12688/f1000research.25088.2

Abstract

Background: In this study we measured the affective appraisal of sounds and video clips using a newly developed graphical self-report tool: the EmojiGrid. The EmojiGrid is a square grid, labeled with emoji that express different degrees of valence and arousal. Users rate the valence and arousal of a given stimulus by simply clicking on the grid. Methods: In Experiment I, observers (N=150, 74 males, mean age=25.2±3.5) used the EmojiGrid to rate their affective appraisal of 77 validated sound clips from nine different semantic categories, covering a large area of the affective space. In Experiment II, observers (N=60, 32 males, mean age=24.5±3.3) used the EmojiGrid to rate their affective appraisal of 50 validated film fragments varying in positive and negative affect (20 positive, 20 negative, 10 neutral). **Results:** The results of this study show that for both sound and video, the agreement between the mean ratings obtained with the EmojiGrid and those obtained with an alternative and validated affective rating tool in previous studies in the literature, is excellent for valence and good for arousal. Our results also show the typical universal U-shaped relation between mean valence and arousal that is commonly observed for affective sensory stimuli, both for sound and video. **Conclusions:** We conclude that the EmojiGrid can be used as an affective self-report tool for the assessment of sound and videoevoked emotions.

Keywords

affective response, audio clips, video clips, EmojiGrid, valence, arousal



Open Peer Review

1. Wei Ming Jonathan Phan, California State University, Long Beach, Long Beach, USA

2. Linda K Kaye D, Edge Hill University, Ormskirk, UK

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Alexander Toet (lextoet@gmail.com)

Author roles: Toet A: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **van Erp JBF**: Funding Acquisition, Methodology, Resources, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2021 Toet A and van Erp JBF. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Toet A and van Erp JBF. Affective rating of audio and video clips using the EmojiGrid [version 2; peer review: 2 approved] F1000Research 2021, 9:970 https://doi.org/10.12688/f1000research.25088.2

First published: 11 Aug 2020, 9:970 https://doi.org/10.12688/f1000research.25088.1

REVISED Amendments from Version 1

We added a concise review of the literature about the emotional affordances of emoji to the Introduction section. In the Data Analysis section, we now explain how the EmojiGrid data were scaled. The graphs in the Results section now represent datapoints by the identifiers of the corresponding stimuli, to allow the visual assessment, comparison and verification of the emotions induced by the different affective stimuli. We also added correlation plots for the mean valence and arousal ratings obtained both with the SAM and EmojiGrid to enable a direct comparison within both of these affective dimensions. In addition, we uploaded a new set of Excel notebooks to the Open Science Framework that include all graphs, together with a brief description of the nature and content of all stimuli, their original affective classification, and their mean valence and arousal values (1) as provided by the authors of the (sound and video) databases and (2) as measured in this study. We extended the Discussion section with some limitations of this study, such as ways to measure mixed emotions, and the fact that the comparison of the SAM and EmojiGrid ratings were based on ratings from different populations.

Any further responses from the reviewers can be found at the end of the article

Introduction

In daily human life, visual and auditory input from our environment significantly determines our feelings, behavior and evaluations (Fazio, 2001; Jaquet et al., 2014; Turley & Milliman, 2000, for a review see: Schreuder et al., 2016). The assessment of the affective response of users to the auditory and visual characteristics of for instance (built and natural) environments (Anderson et al., 1983; Huang et al., 2014; Kuijsters et al., 2015; Ma & Thompson, 2015; Medvedev et al., 2015; Toet et al., 2016; Watts & Pheasant, 2015) and their virtual representations (Houtkamp & Junger, 2010; Houtkamp et al., 2008; Rohrmann & Bishop, 2002; Toet et al., 2013; Westerdahl et al., 2006), multimedia content (Baveye et al., 2018; Soleymani et al., 2015), human-computer interaction systems (Fagerberg et al., 2004; Hudlicka, 2003; Jaimes & Sebe, 2010; Peter & Herbon, 2006; Pfister et al., 2011) and (serious) games (Anolli et al., 2010; Ekman & Lankoski, 2009; Garner et al., 2010; Geslin et al., 2016; Tsukamoto et al., 2010; Wolfson & Case, 2000) is an essential part of their design and evaluation and requires efficient methods to assess whether the desired experiences are indeed achieved. A wide range of physiological, behavioral and cognitive measures is currently available to measure the affective response to sensorial stimuli, each with their own advantages and disadvantages (for a review see: Kaneko et al., 2018a). The most practical and widely used instruments to measure affective responses are questionnaires and rating scales. However, their application is typically time-consuming and requires a significant amount of mental effort (people typically find it difficult to name their emotions, especially mixed or complex ones), which affects the experience itself (Constantinou et al., 2014; Lieberman, 2019; Lieberman et al., 2011; Taylor et al., 2003; Thomassin et al., 2012; for a review see: Torre & Lieberman, 2018) and restricts repeated application. While verbal rating scales are typically more efficient than questionnaires, they also require mental effort since users are required to relate their affective state to verbal descriptions (labels). Graphical rating tools however allow users to intuitively project their feelings to figural elements that correspond to their current affective state.

Arousal and pleasantness (valence) are principal dimensions of affective responses to environmental stimuli (Mehrabian & Russell, 1974). A popular graphical affective self-report tool is the Self-Assessment Mannikin (SAM) (Bradley & Lang, 1994): a set of iconic humanoid figures representing different degrees of valence, arousal, and dominance. Users respond by selecting from each of the three scales the figure that best expresses their own feeling. The SAM has previously been used for the affective rating of video fragments (e.g., Bos et al., 2013; Deng et al., 2017; Detenber et al., 2000; Detenber et al., 1998; Ellard et al., 2012; Ellis & Simons, 2005; Fernández et al., 2012; Soleymani et al., 2008) and auditory stimuli (e.g., Bergman et al., 2009; Bradley & Lang, 2000; Lemaitre et al., 2012; Morris & Boone, 1998; Redondo et al., 2008; Vastfjall et al., 2012). Although the SAM is validated and widely used, users often misunderstand the depicted emotions (Hayashi et al., 2016; Yusoff et al., 2013): especially the arousal dimension (shown as an 'explosion' in the belly area) is often interpreted incorrectly (Betella & Verschure, 2016; Broekens & Brinkman, 2013; Chen et al., 2018; Toet et al., 2018). The SAM also requires a successive assessment of the stimulus on each of its individual dimensions. To overcome these problems we developed an alternative intuitive graphical self-report tool to measure valence and arousal: the EmojiGrid (Toet et al., 2018). The EmojiGrid is a square grid (resembling the Affect Grid: Russell et al., 1989), labeled with emoji that express various degrees of valence and arousal. Emoji are facial icons that can elicit the same range of neural (Gantiva et al., 2020) and emotional (Moore et al., 2013) responses as real human faces. In contrast to photographs, emoji are not associated with overgeneralization (the misattribution of emotions and traits to neutral human faces that merely bear a subtle structural resemblance to emotional expressions: Said et al., 2009), or racial, cultural and sexual biases. Although some facial emoji can be poly-interpretable (Miller et al., 2016; Tigwell & Flatla, 2016) it has been found that emoji with similar facial expressions are typically attributed similar meanings (Jaeger & Ares, 2017; Moore et al., 2013) that are also to a large extent language independent (Novak et al., 2015). Emoji have a wide range of different applications, amongst others in psychological research (Bai et al., 2019). Emoji based rating tools are increasingly becoming popular tools as self-report instruments (Kaye et al., 2017) to measure for instance user and consumer experience (e.g. www.emojiscore. com). Since facial expressions can communicate a wide variety of both basic and complex emotions emoji-based self-report tools may also afford the measurement and expression of mixed (complex) emotions that are otherwise hard to verbalize (Elder, 2018). However, while facial images and emoji are processed in a largely equivalent manner, suggesting that some non-verbal aspects of emoji are processed automatically, further research is required to establish whether they are also emotionally appraised on an implicit level (Kaye et al., 2021).

The EmojiGrid enables users to rate the valence and arousal of a given stimulus by simply clicking on the grid. It has been found that the use of emoji as scale anchors facilitates affective over cognitive responses (Phan *et al.*, 2019). Previous studies on the assessment of affective responses to food images (Toet *et al.*, 2018) and odorants (Toet *et al.*, 2019) showed that the EmojiGrid is self-explaining: valence and arousal ratings did not depend on framing and verbal instructions (Kaneko *et al.*, 2019; Toet *et al.*, 2018). The current study was performed to investigate the EmojiGrid for the affective appraisal of auditory and visual stimuli.

Sounds can induce a wide range of affective and physiological responses (Bradley & Lang, 2000; Gomez & Danuser, 2004; Redondo et al., 2008). Ecological sounds have a clear association with objects or events. However, music can also elicit emotional responses that are as vivid and intense as emotions that are elicited by real-world events (Altenmüller et al., 2002; Gabrielsson & Lindström, 2003; Krumhansl, 1997) and can activate brain regions associated with reward, motivation, pleasure and the mediation of dopaminergic levels (Blood & Zatorre, 2001; Brown et al., 2004; Menon & Levitin, 2005; Small et al., 2001). Even abstract or highly simplified sounds can convey different emotions (Mion et al., 2010; Vastfjall et al., 2012) and can elicit vivid affective mental images when they have some salient acoustic properties in common with the actual sounds. As a result, auditory perception is emotionally biased (Tajadura-Jiménez et al., 2010; Tajadura-Jiménez & Västfjäll, 2008). Video clips can also effectively evoke various affective and physiological responses (Aguado et al., 2018; Carvalho et al., 2012; Rottenberg et al., 2007; Schaefer et al., 2010). While sounds and imagery individually elicit various affective responses that recruit similar brain structures (Gerdes et al., 2014), a wide range of non-linear interactions at multiple processing levels in the brain make that their combined effects are not a priori evident (e.g., Spreckelmeyer et al., 2006; for a review see: Schreuder et al., 2016). Several standardized and validated affective databases have been presented to enable a systematic investigation of sound (Bradley & Lang, 1999; Yang et al., 2018) and video (Aguado et al., 2018; Carvalho et al., 2012; Hewig et al., 2005; Schaefer et al., 2010) elicited affective responses.

This study evaluates the EmojiGrid as a self-report tool for the affective appraisal of auditory and visual events. In two experiments, participants were presented with different sound and video clips, covering both a large part of the valence scale and a wide range of semantic categories. The video clips were stripped of their sound channel (silent) to avoid interaction effects. After perceiving each stimulus, participants reported their affective appraisal (valence and arousal) using the EmojiGrid. The sound samples (Yang *et al.*, 2018) and video clips (Aguado *et al.*, 2018) had been validated in previous studies in the literature using 9-point SAM affective rating scales. This enables an evaluation of the EmojiGrid by directly comparing the mean affective ratings obtained with it to those that were obtained with the SAM.

In this study we also investigate how the mean valence and arousal ratings for the different stimuli are related. Although the relation between valence and arousal for affective stimuli varies between individuals and cultures (Kuppens *et al.*, 2017), it typically shows a quadratic (U-shaped) form across participants (i.e., at the group level): stimuli that are on average rated either high or low on valence are typically also rated as more arousing than stimuli that are on average rated near neutral on valence (Kuppens *et al.*, 2013; Mattek *et al.*, 2017). For the valence and arousal ratings obtained with the EmojiGrid, we therefore also investigate to what extent a quadratic form describes their relation at the group level.

Methods

Participants

English speaking participants from the UK were recruited via the Prolific database (https://www.prolific.co/). Exclusion criteria were age (outside the range of 18–35 years old) and hearing or (color) vision deficiencies. No further attempts were made to eliminate any sampling bias.

We estimated the sample size required for this study with the "ICC.Sample.Size" R-package, assuming an ICC of 0.70 (generally considered as 'moderate': Landis & Koch, 1977), and determined that sample sizes of 57 (Experiment 1) and 23 (Experiment 2) would yield a 95% confidence interval of sufficient precision (± 0.07 ; Landis & Koch, 1977). Because the current experiment was run online and not in a well-controlled laboratory environment, we aimed to recruit about 2–3 times the minimum required number of participants.

This study was approved by the by TNO Ethics Committee (Application nr: 2019-012), and was conducted in accordance with the Helsinki Declaration of 1975, as revised in 2013 (World Medical Association, 2013). Participants electronically signed an informed consent by clicking "*I agree to participate in this study*", affirming that they were at least 18 years old and voluntarily participated in the study. The participants received a small financial compensation for their participation.

Measures

Demographics. The participants in this study reported their nationality, gender and age.

Valence and arousal: the EmojiGrid. The EmojiGrid is a square grid (similar to the Affect Grid: Russell *et al.*, 1989), labeled with emoji that express various degrees of valence and arousal (Figure 1). Users rate their affective appraisal (i.e., the valence and arousal) of a given stimulus by pointing and clicking at the location on the grid that that best represents their impression. The EmojiGrid was originally developed and validated for the affective appraisal of food stimuli, since the SAM appeared to be frequently misunderstood in that context (Toet *et al.*, 2018). It has since also been used and validated for the affective appraisal of doors (Toet *et al.*, 2019).

Procedure

Participants took part in two anonymous online surveys, created with the Gorilla experiment builder (Anwyl-Irvine *et al.*, 2019). After thanking the participants for their interest, the surveys first gave a general introduction to the experiment.

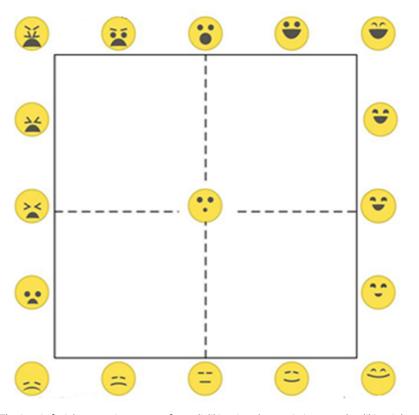


Figure 1. The EmojiGrid. The iconic facial expressions range from disliking (unpleasant) via neutral to liking (pleasant) along the horizontal (valence) axis, while their intensity increases along the vertical (arousal) axis. This figure has been reproduced with permission from Toet *et al.*, 2018.

The instructions asked the participants to perform the survey on a computer or tablet (but not on a device with a small screen such as a smartphone) and to activate the full-screen mode of their browser. This served to maximize the resolution of the questionnaire and to prevent distractions by other programs running in the background. In Experiment I (sounds) the participants were asked to turn off any potentially disturbing sound sources in their room. Then the participants were informed that they would be presented with a given number of different stimuli (sounds in Experiment I and video clips in Experiment II) during the experiment and they were asked to rate their affective appraisal of each stimulus. The instructions also mentioned that it was important to respond seriously, while there would be no correct or incorrect answers. Participants could electronically sign an informed consent. By clicking " I agree to participate in this study ", they confirmed that they were at least 18 years old and that their participation was voluntary. The survey then continued with an assessment of the demographic variables (nationality, gender, age).

Next, the participants were familiarized with the EmojiGrid. First, it was explained how the tool could be used to rate valence and arousal for each stimulus. The instructions were: "To respond, first place the cursor inside the grid on a position that best represents how you feel about the stimulus, and then click the mouse button." Note that the dimensions of valence and arousal were not mentioned here. Then the participants performed two practice trials. In Experiment I, these practice trials also allowed the repeated playing of the sound stimulus. This was done to allow the participants to adjust the sound level of their computer system. The actual experiment started immediately after the practice trials. The stimuli were presented in random order. The participants rated each stimulus by clicking at the appropriate location on the EmojiGrid. The next stimulus appeared immediately after clicking. There were no time restrictions. On average, each experiment lasted about 15 minutes.

Experiment I: Sounds

This experiment served to validate the EmojiGrid as a rating tool for the affective appraisal of sound-evoked emotions. Thereto, participants rated valence and arousal for a selection of sounds from a validated sound database using the EmojiGrid. The results are compared with the corresponding SAM ratings provided for each sound in the database.

Stimuli. The sound stimuli used in this experiment are 77 sound clips from the expanded version of the validated International Affective Digitized Sounds database (IADS-E, available upon request; Yang *et al.*, 2018). The sound clips were selected from 9 different semantic categories: scenarios (2),

breaking sounds (8), daily routine sounds (8), electric sounds (8), people (8), sound effects (8), transport (8), animals (9), and music (10). For all sounds, Yang et al. (2018) provided normative ratings for valence and arousal, obtained with 9-point SAM scales and collected by at least 22 participants from a total pool of 207 young Japanese adults (103 males, 104 females, mean age 21.3 years, SD=2.4). The selection used in the current study was such that the mean affective (valence and arousal) ratings provided for stimuli in the same semantic category were maximally distributed over the two-dimensional affective space (ranging from very negative like a car horn, hurricane sounds or sounds of vomiting, via neutral like people walking up a stairs, to very positive music). As a result, the entire stimulus set is a representative cross-section of the IADS-E covering a large area of the affective space. All sound clips had a fixed duration of 6s. The exact composition of the stimulus set is provided in the Supplementary Material. Each participant rated all sound clips.

Participants. A total of 150 participants (74 males, 76 females) participated in this experiment. All participants were UK nationals. Their mean age was 25.2 (SD= 3.5) years.

Experiment II: Video clips

This experiment served to validate the EmojiGrid as a self-report tool for the assessment of emotions evoke by (silent) video clips. Participants rated valence and arousal for a selection of video clips from a validated set of video fragments using the EmojiGrid. The results are compared with the corresponding SAM ratings for the video clips (Aguado *et al.*, 2018).

Stimuli. The stimuli comprised of a set of 50 film fragments with different affective content (20 positive ones like a coral reef with swimming fishes and jumping dolphins, 10 neutral ones like a man walking in the street or an elevator going down, and 20 negative ones like someone being attacked or a car accident scene). All video clips had a fixed duration of 10 s and were stripped of their soundtracks (for detailed information about the video clips and their availability see Aguado *et al.*, 2018). Aguado *et al.* (2018) obtained normative ratings for valence and arousal, collected by 38 young adults (19 males, 19 females, mean age 22.3 years, SD=2.2) using 9-point SAM scales. In the present study, each participant rated all video clips using the EmojiGrid.

Participants. A total of 60 participants (32 males, 28 females) participated in this experiment. All participants were UK . Their mean age was 24.5 (SD= 3.3) years.

Data analysis

The response data (i.e., the horizontal or valence and vertical or arousal coordinates of the check marks on the EmojiGrid) were quantified as integers between 0 and 550 (the size of the square EmojiGrid in pixels), and then scaled between 1 and 9 for comparison with the results of Yang *et al.* (2018) obtained with a 9-point SAM scale (Experiment I), or between 0 and 8 for comparison with the results of Aguado *et al.* (2018), also obtained with a 9-point SAM scale (Experiment II).

All statistical analyses were performed with IBM SPSS Statistics 26 (www.ibm.com) for Windows. The computation of the intraclass correlation coefficient (ICC) estimates with their associated 95% confidence intervals was based on a mean-rating (k = 3), consistency, 2-way mixed-effects model (Koo & Li, 2016; Shrout & Fleiss, 1979). ICC values less than 0.5 indicate poor reliability, values between 0.5 and 0.75 suggest moderate reliability, values between 0.75 and 0.9 represent good reliability, while values greater than 0.9 indicate excellent reliability (Koo & Li, 2016; Landis & Koch, 1977). For all other analyses a probability level of p < 0.05 was considered to be statistically significant.

MATLAB 2020a was used to further investigate the data. The mean valence and arousal responses were computed across all participants and for each of the stimuli. MATLAB's Curve Fitting Toolbox (version 3.5.7) was used to compute least-squares fits to the data points. Adjusted R-squared values were calculated to quantify the agreement between the data and the curve fits.

Results

Experiment I

Figure 2 shows the correlation plots between the mean valence and arousal ratings for the 77 affective IADS-E sounds used in the current study, obtained with the EmojiGrid (this study) and with a 9-point SAM scale (Yang *et al.* (2018). This figure illustrates the overall agreement between the affective ratings obtained with both self-assessment tools for affective sound stimuli.

The linear (two-tailed) Pearson correlation coefficients between the valence and arousal ratings obtained with the EmojiGrid (present study) and with the SAM (Yang *et al.*, 2018) were, respectively, 0.881 and 0.760 (p<0.001). To further quantify the agreement between both rating tools we computed intraclass correlation coefficients (ICC) with their 95% confidence intervals for the mean valence and arousal ratings between both studies. The ICC value for valence is 0.936 [0.899–0.959] while the ICC for arousal is 0.793 [0.674–0.868], indicating both studies show an excellent agreement for valence and a good agreement for arousal (even though the current study was performed via the internet and therefore did not provide the amount of control over many experimental factors as one would have in a lab experiment).

Figure 3 shows the relation between the mean valence and arousal ratings for the 77 IADS-E sounds used as stimuli in the current study, measured both with the EmojiGrid (this study) and with a 9-point SAM scale (Yang *et al.* (2018). The curves in this figure represent least-squares quadratic fits to the data points. The adjusted R-squared values are 0.62 for results obtained with the EmojiGrid and 0.22 for the SAM results. Hence, both

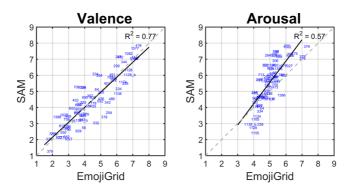


Figure 2. Relation between mean valence (left) and arousal (right) ratings obtained with the SAM and EmojiGrid for selected sounds from the IADS-E database. Labels correspond to the original identifiers of the stimuli (Yang *et al.*, 2018). The line segments represent linear fits to the data points.

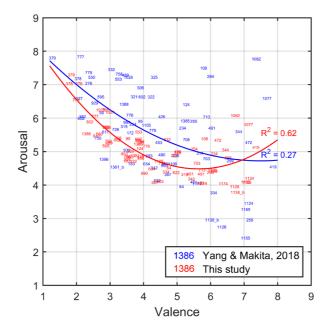


Figure 3. Relation between mean valence and arousal ratings for selected sounds from the IADS-E database. Labels correspond to the original identifiers of the stimuli (Yang *et al.*, 2018). Blue labels represent data obtained with the SAM (Yang *et al.*, 2018), while red labels represent data obtained with the EmojiGrid (this study). The curves represent quadratic fits to the corresponding data points.

methods yield a relation between mean valence and arousal ratings that can indeed be described by a quadratic (U-shaped) relation at the nomothetic (group) level.

Experiment II

Figure 4 shows the correlation plots between the mean valence and arousal ratings for the 50 affective video clips used in the current study, obtained with the EmojiGrid (this study) and with a 9-point SAM scale (Aguado *et al.*, 2018). This figure illustrates the overall agreement between the affective ratings obtained with both self-assessment tools for affective sound stimuli.

The linear (two-tailed) Pearson correlation coefficients between the valence and arousal ratings obtained with the EmojiGrid (present study) and with the SAM (Aguado *et al.*, 2018) were respectively 0.963 and 0.624 (p<0.001). To further quantify the agreement between both rating tools we computed intraclass correlation coefficients (ICC) with their 95% confidence intervals for the mean valence and arousal ratings between both studies. The ICC value for valence is 0.981 [0.967 – 0.989] while the ICC for arousal is 0.721 [0.509 – 0.842], indicating both studies show an excellent agreement for valence and a good agreement for arousal.

Figure 5 shows the relation between the mean valence and arousal ratings for the 50 video clips tested. The curves in this figure represent quadratic fits to the data points. The adjusted R-squared values are respectively 0.68 and 0.78. Hence, both methods yield a relation between mean valence and arousal

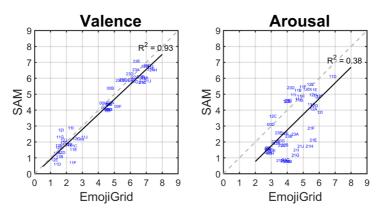


Figure 4. Relation between mean valence (left) and arousal (right) ratings obtained with the SAM and EmojiGrid for 50 affective video clips (Aguado *et al.*, 2018). Labels correspond to the original identifiers of the stimuli (Yang *et al.*, 2018). The line segments represent linear fits to the data points.

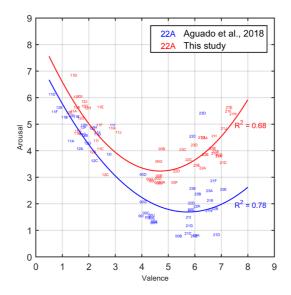


Figure 5. Mean valence and arousal ratings for affective film clips. Labels correspond to the original identifiers of the stimuli (Aguado *et al.*, 2018). Blue labels represent data obtained with the SAM (Aguado *et al.*, 2018) while red labels represent data obtained with the EmojiGrid (this study). The curves show quadratic fits to the corresponding data points.

ratings that can be described by a quadratic (U-shaped) relation at the nomothetic (group) level.

Raw data from each experiment are available as *Underlying data* (Toet, 2020).

Conclusion

In this study we evaluated the recently developed EmojiGrid self-report tool for the affective rating of sounds and video. In two experiments, observers rated their affective appraisal of sound and video clips using the EmojiGrid. The results show a close correspondence between the mean ratings obtained with the EmojiGrid and those obtained with the validated

SAM tool in previous validation studies in the literature: the agreement is excellent for valence and good for arousal, both for sound and video. Also, for both sound and video, the EmojiGrid yields the universal U-shaped (quadratic) relation between mean valence and arousal that is typically observed for affective sensory stimuli. We conclude that the EmojiGrid is an efficient affective self-report tool for the assessment of sound and video-evoked emotions.

A limitation of the EmojiGrid is the fact that it is based on the circumplex model of affect which posits that positive and negative feelings are mutually exclusive (Russell, 1980). Hence, in its present form, and similar to other affective self-report tools like the SAM or VAS scales, the EmojiGrid only allows the measurement of a single emotion at a time. However, emotions are not strictly bipolar and two or more same or opposite valenced emotions can co-occur together (Larsen & McGraw, 2014; Larsen et al., 2001). Mixed emotions consisting of opposite feelings can in principle be registered with the EmojiGrid by allowing participants to enter multiple responses.

Another limitation of this study is the fact that the comparison of the SAM and EmojiGrid ratings were based on ratings from different populations (akin to a comparison of two independent samples). Hence, our current regression estimates are optimized based on the particular samples that were used. Future studies should investigate a design in which the same participants use both self-report tools to rate the same set of stimuli.

Future applications of the EmojiGrid may involve the real-time evaluation of affective events or the provision of affective feedback. For instance, in studies on affective communication in human-computer interaction (e.g., Tajadura-Jiménez & Västfjäll, 2008), the EmojiGrid can be deployed as a continuous response tool by moving a mouse-controlled cursor over the grid while logging the cursor coordinates. Such an implementation may also afford the affective annotation of multimedia (Chen et al., 2007; Runge et al., 2016), and could be useful for personalized affective video retrieval or recommender systems (Hanjalic & Xu, 2005; Koelstra et al., 2012; Lopatovska & Arapakis, 2011; Xu et al., 2008), for real-time affective appraisal of entertainment (Fleureau et al., 2012) or to provide affective input to serious gaming applications (Anolli et al., 2010) and affective music generation (Kim & André, 2004). Sensiks (www.sensiks.com) has adopted a simplified version of the EmojiGrid in its Sensory Reality Pod to enable the user to select and tune multisensory (visual, auditory, tactile and olfactory) affective experiences.

Data availability

Underlying data

Open Science Framework: Affective rating of audio and video clips using the EmojiGrid. https://doi.org/10.17605/OSF.IO/ GTZH4 (Toet, 2020).

File 'Results_sound_video' (XLSX) contains the EmojiGrid co-ordinates selected by each participant following each stimulus.

Open Science Framework: Additional data on affective rating of audio and video clips using the EmojiGrid. https://doi. org/10.17605/OSF.IO/6HQTR

File 'sound_results.xlsx' contains the mean valence and arousal ratings, obtained with the SAM (Yang et al., 2018) and the EmojiGrid (this study), together with graphs in which each of the stimuli are labelled for easy identification.

File 'video_results.xlsx' contains the mean valence and arousal ratings, obtained with the SAM (Aguado et al., 2018) and the EmojiGrid (this study), together with graphs in which each of the stimuli are labelled for easy identification.

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Acknowledgements

The authors thank Dr. Wanlu Yang (Hiroshima University, Higashi-Hiroshima, Japan) for providing the IADS-E sound database, and Dr. Luis Aguado (Universidad Complutense de Madrid, Spain) for providing the validated movie clips.

References

Aquado L. Fernández-Cahill M. Román FL et al.: Fvaluative and psychophysiological responses to short film clips of different emotional content. J Psychophysiol. 2018; 32(1): 1-19.

Altenmüller E, Schürmann K, Lim VK, et al.: Hits to the left, flops to the right: different emotions during listening to music are reflected in cortical lateralisation patterns. Neuropsychologia. 2002; 40(13): 2242-2256. PubMed Abstract | Publisher Full Text

Anderson LM, Mulligan BE, Goodman LS, et al.: Effects of sounds on preferences for outdoor settings. Environ Behav. 1983; 15(5): 539-566. Publisher Full Text

Anolli L, Mantovani F, Confalonieri L, et al.: Emotions in serious games: From experience to assessment. International Journal of Emerging Technologies in Learning, 2010: 5(Special Issue 2): 7-16.

Reference Source

Anwyl-Irvine A. Massonnié I. Flitton A. et al.: Gorilla in our Midst: An online behavioral experiment builder. bioRxiv. 2019; 438242. Publisher Full Text

Bai Q, Dan Q, Mu Z, et al.: A systematic review of Emoji: Current research and future perspectives. Front Psychol. 2019; 10: 2221. PubMed Abstract | Publisher Full Text | Free Full Text

Baveye Y, Chamaret C, Dellandréa E, et al.: Affective video content analysis: a multidisciplinary insight. IEEE Trans Affect Comput. 2018; 9(4): 396-409. **Publisher Full Text**

Bergman P, Sköld A, Västfjäll D, et al.: Perceptual and emotional categorization of sound. J Acoust Soc Am. 2009; 126(6): 3156-3167. PubMed Abstract | Publisher Full Text

Betella A, Verschure PFMJ: The Affective Slider: A digital self-assessment scale for the measurement of human emotions. PLoS One. 2016; 11(2): e0148037

PubMed Abstract | Publisher Full Text | Free Full Text

Blood AJ, Zatorre RJ: Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. Proc Natl Acad Sci U S Á. 2001; 98(20): 11818–11823. PubMed Abstract | Publisher Full Text | Free Full Text

Bos MGN, Jentgens P, Beckers T, et al.: Psychophysiological response patterns to affective film stimuli. PLoS One. 2013; 8(4): e62661. PubMed Abstract | Publisher Full Text | Free Full Text

Bradley MM, Lang PJ: International affective digitized sounds (IADS):

Publisher Full Text

Bradley MM, Lang PJ: Measuring emotion: the Self-Assessment Manikin and the semantic differential. J Behav Ther Exp Psychiatry. 1994; 25(1): 49-59. PubMed Abstract | Publisher Full Text

Stimuli, instruction manual and affective ratings. (Gainesville, FL: The Center for Research in Psychophysiology, University of Florida). 1999. **Reference Source**

Bradley MM, Lang PJ: Affective reactions to acoustic stimuli. Psychophysiology. 2000; **37**(2): 204–215.

PubMed Abstract | Publisher Full Text

Broekens J, Brinkman WP: AffectButton: A method for reliable and valid affective self-report. Int J Hum Comput Stud. 2013; 71(6): 641-667. **Publisher Full Text**

Brown S, Martinez MJ, Parsons LM: Passive music listening spontaneously engages limbic and paralimbic systems. Neuroreport. 2004; 15(13): 2033-2037

PubMed Abstract | Publisher Full Text

Carvalho S, Leite J, Galdo-Álvarez S, *et al.*: **The emotional movie database** (EMDB): A self-report and psychophysiological study. *Appl Psychophysiol Biofeedback*. 2012; **37**(4): 279–294. PubMed Abstract | Publisher Full Text

Chen L, Chen GC, Xu CZ, et al.: EmoPlayer: A media player for video clips with affective annotations. Interact Comput. 2007; 20(1): 17-28. **Publisher Full Text**

Chen Y, Gao Q, Lv Q, et al.: Comparing measurements for emotion evoked by oral care products. Int J Ind Ergon. 2018; 66: 119-129. **Publisher Full Text**

Constantinou E, Van Den Houte M, Bogaerts K, et al.: Can words heal? Using affect labeling to reduce the effects of unpleasant cues on symptom reporting. Front Psychol. 2014; 5: 807. PubMed Abstract | Publisher Full Text | Free Full Text

Deng Y, Yang M, Zhou R: A new standardized emotional film database for asian culture. Front Psychol. 2017; 8: 1941. PubMed Abstract | Publisher Full Text | Free Full Text

Detenber BH, Simons RF, Reiss JE: The emotional significance of color in television presentations. Media Psychol. 2000; 2(4): 331–355. **Publisher Full Text**

Detenber BH, Simons RF, Bennett GG Jr: Roll 'em!: The effects of picture motion on emotional responses. J Broadcast Electron Media. 1998; 42(1): 113-128.

Publisher Full Text

Ekman I, Lankoski P: Hair-raising entertainment: Emotions, sound, and structure in Silent Hill 2 and Fatal Frame. In: Horro video games. Essyas on the fusion of fear and play. ed. B. Perron. Jefferson, NC USA: Mc Farland & Company, Inc., 2009; 181–199. Reference Source

Elder AM: What words can't say: Emoji and other non-verbal elements of technologically-mediated communication. Journal of Information, Communication and Ethics in Society. 2018; 16(1): 2–15.

Publisher Full Text

Ellard KK, Farchione TJ, Barlow DH: Relative effectiveness of emotion induction procedures and the role of personal relevance in a clinical sample: A comparison of film, images, and music. J Psychopathol Behav Assess. 2012; 34(2): 232-243. **Publisher Full Text**

Ellis RJ, Simons RF: **The impact of music on subjective and physiological indices of emotion while viewing films**. *Psychomusicology: A Journal of* Research in Music Cognition. 2005; 19(1): 15-40. **Publisher Full Text**

Fagerberg P, Ståhl A, Höök K: eMoto: emotionally engaging interaction. Pers Ubiquitous Comput. 2004; 8(5): 377-381.

Publisher Full Text

Fazio RH: On the automatic activation of associated evaluations: An overview. Cognition & Emotion. 2001; 15(2): 115-141. **Publisher Full Text**

Fernández C, Pascual J, Soler J, et al.: Physiological responses induced by emotion-eliciting films. Appl Psychophysiol Biofeedback. 2012; 37(2): 73-79. PubMed Abstract | Publisher Full Text

Fleureau J, Guillotel P, Quan H: Physiological-based affect event detector for entertainment video applications. IEEE Trans Affect Comput. 2012; 3(3): 379-385.

Publisher Full Text

Gabrielsson A, Lindström Wik S: Strong experiences related to music: A descriptive system. Music Sci. 2003; 7(2): 157-217. Publisher Full Text

Gantiva C, Sotaquirá M, Araujo A, et al.: Cortical processing of human and emoji faces: an ERP analysis. Behaviour & Information Technology. 2020; 39(8): 935-943.

Publisher Full Text

Garner T, Grimshaw M, Abdel Nabi D: A preliminary experiment to assess the fear value of preselected sound parameters in a survival horror game. 5th Audio Mostly Conference: A Conference on Interaction with Sound (AM'10). New York, NY USA: ÁCM. 2010; 1–9. **Publisher Full Text**

Gerdes ABM, Wieser MJ, Alpers GW: Emotional pictures and sounds: A review of multimodal interactions of emotion cues in multiple domains. Front Psychol. 2014; 5: 1351.

PubMed Abstract | Publisher Full Text | Free Full Text

Geslin E, Jégou L, Beaudoin D: **How color properties can be used to elicit emotions in video games.** *International Journal of Computer Games Technology.* 2016; **2016**(Article ID 5182768): 1–9. **Publisher Full Text**

Gomez P, Danuser B: Affective and physiological responses to environmental noises and music. Int J Psychophysiol. 2004; 53(2): 91–103. PubMed Abstract | Publisher Full Text

Hanjalic A, Xu LQ: Affective video content representation and modeling. IEEE Trans Multimedia. 2005; 7(1): 143-154. **Publisher Full Text**

Hayashi ECS, Gutiérrez Posada JE, Maike VRML, et al.: Exploring new formats of the Self-Assessment Manikin in the design with children. 15th Brazilian Symposium on Human Factors in Computer Systems. New York, NY USA: ACM. 2016: 1-10.

Publisher Full Text

Hewig J, Hagemann D, Seifert J, et al.: A revised film set for the induction of basic emotions. Cognition & Emotion. 2005; 19(7): 1095-1109. **Publisher Full Text**

Houtkamp JM, Junger MLA: Affective qualities of an urban environment on a Hotstanp Jm, Janger Marchand Conference Information Visualisation., ed. E. Banissi, Los Alamitos, CA USA: IEEE Computer Society. 2010; 597–603. Publisher Full Text

Houtkamp JM, Schuurink EL, Toet A: Thunderstorms in my computer: the effect of visual dynamics and sound in a 3D environment. eds. M Bannatyne & J Counsell: IEEE Computer Society, 2008; 11–17. **Publisher Full Text**

Huang H, Klettner S, Schmidt M, et al.: AffectRoute - considering people's affective responses to environments for enhancing route-planning services. Int J Geogr Inf Sci. 2014; 28(12): 2456–2473. **Publisher Full Text**

Hudlicka E: To feel or not to feel: the role of affect in human-computer interaction. Int J Hum Comput Stud. 2003; 59(1–2): 1–32. Publisher Full Text

Jaeger SR, Ares G: Dominant meanings of facial emoji: Insights from Chinese consumers and comparison with meanings from internet resources. Food Quality and Preference. 2017; 62: 275–283. **Publisher Full Text**

Jaimes A, Sebe N: Multimodal human-computer interaction: a survey. Comput Vis Image Underst. 2010; 108(1-2): 116-134. **Publisher Full Text**

Jaquet L, Danuser B, Gomez P: Music and felt emotions: How systematic pitch level variations affect the experience of pleasantness and arousal. Psychol Music. 2014; 42(1): 51-70. **Publisher Full Text**

Kaneko D, Toet A, Brouwer AM, et al.: Methods for evaluating emotions evoked by food experiences: A literature review. Front Psychol. 2018a; 9: 911. PubMed Abstract | Publisher Full Text | Free Full Text

Kaneko D, Toet A, Ushiama S, *et al.*: EmojiGrid: a 2D pictorial scale for cross-cultural emotion assessment of negatively and positively valenced food. Food Res Int. 2019; 115: 541-551.

PubMed Abstract | Publisher Full Text

Kaye LK, Malone SA, Wall HJ: Emojis: Insights, affordances, and possibilities for psychological science. Trends Cogn Sci. 2017; 21(2): 66-68. PubMed Abstract | Publisher Full Text

Kaye LK, Rodriguez-Cuadrado S, Malone SA, et al.: How emotional are emoji?: Exploring the effect of emotional valence on the processing of emoji stimuli. Computers in Human Behavior. 2021; 116: 106648. Publisher Full Text

Kim S, André E: Composing affective music with a generate and sense approach. In: Flairs 2004 - Special Track on AI and Music. AAAI. 2004. **Reference Source**

Koelstra S, Muhl C, Soleymani M, et al.: DEAP: A database for emotion analysis using physiological signals. IEEE Trans Affect Comput. 2012; 3(1): 18-31.

Publisher Full Text

Koo TK, Li MY: A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016; 15(2): 155-163. PubMed Abstract | Publisher Full Text | Free Full Text

Krumhansl CL: An exploratory study of musical emotions and psychophysiology. Can J Exp Psychol. 1997; 51(4): 336-353. PubMed Abstract | Publisher Full Text

Kuijsters A, Redi J, de Ruyter B, et al.: Affective ambiences created with lighting for older people. Light Res Technol. 2015; 47(7): 859-875. **Publisher Full Text**

Kuppens P, Tuerlinckx F, Russell JA, et al.: The relation between valence and arousal in subjective experience. *Psychol Bull.* 2013; **139**(4): 917–940. PubMed Abstract | Publisher Full Text

Kuppens P, Tuerlinckx F, Yik M, et al.: The relation between valence and arousal in subjective experience varies with personality and culture. J Pers. 2017; 85(4): 530-542.

PubMed Abstract | Publisher Full Text

Landis JR, Koch GG: The measurement of observer agreement for

categorical data. Biometrics. 1977; 33(1): 159-174. PubMed Abstract | Publisher Full Text

Larsen JT, McGraw AP: The case for mixed emotions. Soc Personal Psychol Compass. 2014; 8(6): 263–274. Publisher Full Text

Larsen JT, McGraw AP, Cacioppo JT: Can people feel happy and sad at the same time? J Pers Soc Psychol. 2001; 81(4): 684–696. PubMed Abstract | Publisher Full Text

Lemaitre G, Houix O, Susini P, et al.: Feelings elicited by auditory feedback from a computationally augmented artifact: The flops. *IEEE Trans Affect Comput.* 2012; **3**(3): 335–348. **Publisher Full Text**

Lieberman MD: Affect labeling in the age of social media. Nat Hum Behav. 2019: 3(1): 20-21.

PubMed Abstract | Publisher Full Text

Lieberman MD, Inagaki TK, Tabibnia G, et al.: Subjective responses to emotional stimuli during labeling, reappraisal, and distraction. Emotion. 2011; 11(3): 468-480.

PubMed Abstract | Publisher Full Text | Free Full Text

Lopatovska I, Arapakis I: Theories, methods and current research on emotions in library and information science, information retrieval and human-computer interaction. Inf Process Manag. 2011; 47(4): 575-592. **Publisher Full Text**

Ma W, Thompson WF: Human emotions track changes in the acoustic environment. Proc Natl Acad Sci U S A. 2015; **112**(47): 14563–14568. PubMed Abstract | Publisher Full Text | Free Full Text

Mattek AM, Wolford GL, Whalen PJ: A mathematical model captures the structure of subjective affect. Perspect Psychol Sci. 2017; **12**(3): 508–526. PubMed Abstract | Publisher Full Text | Free Full Text

Medvedev O, Shepherd D, Hautus MJ: The restorative potential of soundscapes: A physiological investigation. Appl Acoust. 2015; 96: 20-26. Publisher Full Text

Mehrabian A, Russell JA: An approach to environmental psychology. Boston, MA, USA: The MIT Press. 1974 **Reference Source**

Menon V, Levitin DJ: The rewards of music listening: Response and physiological connectivity of the mesolimbic system. Neuroimage. 2005; 28(1): 175–184.

PubMed Abstract | Publisher Full Text

Miller H, Thebault-Spieker J, Chang S, et al.: Blissfully happy" or "ready to fight": Varying Interpretations of Emoji. Paper presented at the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016). AAAI, Menlo Park, CA. 2016; 259–268. **Reference Source**

Mion L, D'Incá G, de Götzen A, et al.: Modeling expression with perceptual audio features to enhance user interaction. Computer Music Journal. 2010; 34(1): 65-79.

Publisher Full Text

Moore A, Steiner CM, Conlan O: Design and development of an empirical smiley-based affective instrument. Paper presented at the 21st Conference on User Modeling, Adaptation, and Personalization,. CEUR Workshop Proceedings, Rome, Italy. 2013; 997: 41-52.

Reference Source

Morris JD, Boone MA: The effects of music on emotional response, brand attitude, and purchase intent in an emotional advertising condition. Adv Consum Res. 1998; 25(1): 518-526. Reference Source

Novak KP, Smailović J, Sluban B, et al.: Sentiment of emojis. PLoS One. 2015; 10(12): e0144296. PubMed Abstract | Publisher Full Text | Free Full Text

Peter C, Herbon A: Emotion representation and physiology assignments in digital systems. Interact Comput. 2006; 18(2): 139-170. **Publisher Full Text**

Pfister HR, Wollstädter S, Peter C: Affective responses to system messages in human-computer-interaction: Effects of modality and message type. Interact Comput. 2011; 23(4): 372–383. **Publisher Full Text**

Phan WMJ, Amrhein R, Rounds J, et al.: Contextualizing interest scales with emojis: Implications for measurement and validity. J Career Assess. 2019; 27(1): 114-133.

Publisher Full Text

Redondo J, Fraga I, Padrón I, et al.: Affective ratings of sound stimuli. Behav Res Methods. 2008; 40(3): 784-790.

PubMed Abstract | Publisher Full Text

Rohrmann B, Bishop ID: Subjective responses to computer simulations of urban environments. J Environ Psychol. 2002; 22(4): 319-331. Publisher Full Text

Rottenberg J, Ray RR, Gross JJ: Emotion elicitation using films. eds. J.A. Coan & J.J.B. Allen: Oxford University Press, 2007; 9-28 **Reference Source**

Runge N, Hellmeier M, Wenig D, et al.: Tag your emotions: a novel mobile

user interface for annotating images with emotions. in: 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct. 2961836: ACM, 2016; 846-853. **Publisher Full Text**

Russell JA: A circumplex model of affect. Journal of Personality and Social Psychology. 1980; **39**(6): 1161-1178. Publisher Full Text

Russell JA, Weiss A, Mendelson GA: Affect grid: A single-item scale of pleasure and arousal. / Pers Soc Psychol. 1989; 57(3): 493-502. Publisher Full Text

Said CP, Sebe N, Todorov A: Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. Emotion. 2009; 9(2): 260-264.

PubMed Abstract | Publisher Full Text

Schaefer A, Nils F, Sanchez X, *et al.*: **Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers.** *Cogn Emot.* 2010; **24**(7): 1153–1172. **Publisher Full Text**

Schreuder E, van Erp J, Toet A, *et al.*: **Emotional responses to multisensory environmental stimuli.** *SAGE Open.* 2016; **6**(1): 1–19. **Publisher Full Text**

Shrout PE, Fleiss JL: Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull.* 1979; 86(2): 420–428. PubMed Abstract | Publisher Full Text

Small DM, Zatorre RJ, Dagher A, et al.: Changes in brain activity related to eating chocolate: from pleasure to aversion. Brain. 2001; 124(Pt 9): 1720-1733.

PubMed Abstract | Publisher Full Text

Soleymani M, Chanel G, Kierkels JJM, et al.: Affective ranking of movie scenes using physiological signals and content analysis. In: 2nd ACM workshop on Multimedia semantics. New York, NY, USA: ACM, 2008; 32–39. **Publisher Full Text**

Soleymani M, Yang Y, Irie G, et al.: Guest editorial: Challenges and perspectives for affective analysis in multimedia. IEEE Trans Affect Comput. 2015; 6(3): 206-208.

Publisher Full Text

Spreckelmeyer KN, Kutas M, Urbach TP, et al.: Combined perception of emotion in pictures and musical sounds. Brain Res. 2006; 1070(1): 160-170. PubMed Abstract | Publisher Full Text

Tajadura-Jiménez A, Väljamäe A, Asutay E, et al.: Embodied auditory perception: the emotional impact of approaching and receding sound sources. Emotion. 2010; 10(2): 216-229.

PubMed Abstract | Publisher Full Text

Tajadura-Jiménez A, Västfjäll D: Auditory-induced emotion: A neglected channel for communication in human-computer interaction. In: Affect and Emotion in Human-Computer Interaction. eds. C. Peter & B. R. Berlin - Heidelberg, Germany: Springer, 2008; 63–74.

Publisher Full Text

Taylor SF, Phan KL, Decker LR, et al.: Subjective rating of emotionally salient stimuli modulates neural activity. NeuroImage. 2003; 18(3): 650–659 **Publisher Full Text**

Thomassin K, Morelen D, Suveg C: **Emotion reporting using electronic diaries** reduces anxiety symptoms in girls with emotion dysregulation. *J Contemp Psychother*. 2012; **42**(4): 207–213.

Publisher Full Text

Tigwell GW, Flatla DR: Oh that's what you meant!: reducing emoji misunderstanding. Paper presented at the Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, ACM 2961844. 2016; 859-866.

Publisher Full Text

Toet A: Affective rating of audio and video clips using the EmojiGrid. 2020. http://www.doi.org/10.17605/OSF.IO/GTZH4

Toet A, Eijsman S, Liu Y, et al.: The relation between valence and arousal in subjective odor experience. Chemosens Percept. Online first. 2020; 13: 141-151.

Publisher Full Text

Toet A, Houtkamp JM, van der Meulen R: Visual and auditory cue effects on risk assessment in a highway training simulation. Simul Games. 2013; 44(5): 732-753

Publisher Full Text

Toet A, Houtkamp JM, Vreugdenhil PE: Effects of personal relevance and simulated darkness on the affective appraisal of a virtual environment. Peerl. 2016: 4: e1743.

PubMed Abstract | Publisher Full Text | Free Full Text

Toet A, Kaneko D, Ushiama S, et al.: EmojiGrid: A 2D pictorial scale for the assessment of food elicited emotions. Front Psychol. 2018; 9: 2396. PubMed Abstract | Publisher Full Text | Free Full Text

Torre JB, Lieberman MD: Putting feelings into words: Affect labeling as implicit emotion regulation. *Emotion Review*. 2018; **10**(2): 116–124. **Publisher Full Text**

Tsukamoto M, Yamada M, Yoneda R: A dimensional study on the emotion

of musical pieces composed for video games. In: 20th International Congress on Acoustics 2010 (ICA 2010). eds. M. Burgess, J. Davey, C. Don & T. McMinn. Australian Acoustical Society, 2010; 4058–4060. **Reference Source**

Turley LW, Milliman RE: Atmospheric effects on shopping behavior: A review of the experimental evidence. J Bus Res. 2000; 49(2): 193-211 **Publisher Full Text**

Vastfjall D, Bergman P, Sköld A, et al.: Emotional responses to information and warning sounds. Journal of Ergonomics. 2012; 2(3): 106 **Publisher Full Text**

Watts GR, Pheasant RJ: Tranquillity in the Scottish Highlands and Dartmoor National Park - The importance of soundscapes and emotional factors. Appl Acoust. 2015; 89: 297-305.

Publisher Full Text

Westerdahl B, Suneson K, Wernemyr C, et al.: Users' evaluation of a virtual reality architectural model compared with the experience of the completed building. *Autom Constr.* 2006; **15**(2): 150–165. Publisher Full Text

Wolfson S, Case G: The effects of sound and colour on responses to a

computer game. Interact Comput. 2000; 13(2): 183-192. **Publisher Full Text**

World Medical Association: World Medical Association declaration of Helsinki: Ethical principles for medical research involving human subjects. JAMA. 2013; **310**(20): 2191–2194. PubMed Abstract | Publisher Full Text

Xu C, Chen L, Chen G: **A color bar based affective annotation method for media player**. In: *Frontiers of WWW Research and Development - APWeb 2006.* eds. X. Zhou, J. Li, H.T. Shen, M. Kitsuregawa & Y. Zhang. Heidelberg/Berlin, Germany: Springer, 2008; 759–764. Publisher Full Text

Yang W, Makita K, Nakao T, et al.: Affective auditory stimulus database: An

expanded version of the International Affective Digitized Sounds (IADS-E). Behav Res Methods. 2018; 50(4): 1415–1429. PubMed Abstract | Publisher Full Text

Yusoff YM, Ruthven I, Landoni M: **Measuring emotion: A new evaluation tool for very young children**. In: *4th Int. Conf. on Computing and Informatics (ICOCI 2013).* Sarawak, Malaysia: Universiti Utara Malaysia, 2013; 358–363. **Reference Source**

Open Peer Review

Current Peer Review Status: 💙

Version 2

Reviewer Report 23 April 2021

https://doi.org/10.5256/f1000research.55394.r82777

© **2021 Phan W.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Wei Ming Jonathan Phan

California State University, Long Beach, Long Beach, CA, USA

Thank you very much the updated paper. The authors has addressed my concerns adequately.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Survey methodology, rating formats, Emotions, and Emojis.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 07 April 2021

https://doi.org/10.5256/f1000research.55394.r82776

© **2021 Kaye L.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Linda K Kaye 匝

Department of Psychology, Edge Hill University, Ormskirk, UK

Thank you to the authors for making relevant changes as noted from the first round of reviews. I am satisfied that these have been addressed sufficiently.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Psychology of emoji; cyberpsychology; online behaviour

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 11 January 2021

https://doi.org/10.5256/f1000research.27685.r76598

© **2021 Kaye L.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

了 🛛 Linda K Kaye 匝

Department of Psychology, Edge Hill University, Ormskirk, UK

This is an interesting study that seeks to validate the EmojiGrid for use with auditory and video stimuli. Thank you to the authors for providing the research resources on OSF as this is helpful when reviewing the research. Overall, the research has merits but would benefit from being more detailed especially in the introductory and discussion sections. I also have a methodological query but this may be rectified from additional clarity in the writing of this section.

1. The introduction could do with additional literature about the emotional affordances of emoji. That is, the research is presented as assuming that emoji are emotional stimuli but does not provide a review of the literature which can support this. Interestingly, recent evidence (Kaye *et al.*, 2021) suggests that emoji may not be processed emotionally on an implicit level, so the authors should be careful about their assumptions in this regard. Relevant sources that may be useful:

Bai, Q., Dan, Q., Mu, Z., & Yang, M. (2019). A systematic review of emoji: Current research and future perspectives. *Frontiers in Psychology, 10*, e2221. doi:10.3389/fpsyg.2019.02221¹ Derks, D., Fischer, A. H., & Bos, A. E. R. (2008). The role of emotion in computer-mediated communication: A review. *Computers in Human Behavior, 24* (3), 766-785² Kaye, L. K., Rodriguez Cuadrado, S., Malone, S. A., Wall, H. J., Gaunt, E., Mulvey, A. L., & Graham, C. (2021). How emotional are emoji?: Exploring the effect of emotional valence on the processing of emoji stimuli. *Computers in Human Behavior, 116*, 106648³ Novak, P. K., Smailović, J., Sluban, B., & Mozetič, I. (2015). Sentiment of emojis. *PLoS ONE, 10* (12), e0144296⁴

2. With regards to the data presented (e.g., Fig 2), it is not made explicitly clear how numerical values were determined based on the responses from the EmojiGrid. E.g., how are each of the emoji symbols based on their position on the axis determined numerically? From Fig 1, it looks like this ranges from 1 to 5 based on the number of emoji on each axis. However, looking in the methodology, the SAM scale is outlined as being a 9-item response scale so it isn't clear how Fig 2 & 3 can present the data from these two scales on the same axis if the response scales are different.

3. The discussion could benefit from further elaboration. E.g., To what extent do the findings contribute theoretically to the literature? What are the limitations of the work?

Minor

1. In the methodology, it is more typical to use the term "participants" rather than "persons"

References

1. Bai Q, Dan Q, Mu Z, Yang M: A Systematic Review of Emoji: Current Research and Future Perspectives. *Frontiers in Psychology*. 2019; **10**. Publisher Full Text

2. Derks D, Fischer A, Bos A: The role of emotion in computer-mediated communication: A review. *Computers in Human Behavior*. 2008; **24** (3): 766-785 Publisher Full Text

3. Kaye L, Rodriguez-Cuadrado S, Malone S, Wall H, et al.: How emotional are emoji?: Exploring the effect of emotional valence on the processing of emoji stimuli. *Computers in Human Behavior*. 2021; **116**. Publisher Full Text

4. Kralj Novak P, Smailović J, Sluban B, Mozetič I: Sentiment of Emojis.*PLoS One*. 2015; **10** (12): e0144296 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature? Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others? Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Psychology of emoji; cyberpsychology; online behaviour

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 10 Mar 2021

Alexander Toet, TNO, Soesterberg, The Netherlands

Dear Dr Kaye,

Thank you for your critical remarks and valuable suggestions which definitely helped us to improve our initial draft paper. Also, we appreciate the fact that you spent your valuable time on this review.

1. Literature about the emotional affordances of emoji

Thank you for this suggestion. We agree that reviewing literature about the emotional affordances of emoji will be a valuable addition to the Introduction, helping the reader to better place the current findings in their context. We therefore added the following text to the Introduction:

"Emoji are facial icons that can elicit the same range of neural (Gantiva, Sotaquirá, Araujo, & Cuervo, 2020) and emotional (Moore, Steiner, & Conlan, 2013) responses as real human faces. In contrast to photographs, emoji are not associated with overgeneralization (the misattribution of emotions and traits to neutral human faces that merely bear a subtle structural resemblance to emotional expressions: Said, Sebe, & Todorov, 2009), or racial, cultural and sexual biases. Although some facial emoji can be poly-interpretable (Miller et al., 2016; Tigwell & Flatla, 2016) it has been found that emoji with similar facial expressions are typically attributed similar meanings (Jaeger & Ares, 2017; Moore et al., 2013) that are also to a large extent language independent (Kralj Novak, Smailović, Sluban, & Mozetič, 2015). Emoji have a wide range of different applications, amongst others in psychological research (Bai, Dan, Mu, & Yang, 2019). Emoji based rating tools are increasingly becoming popular tools as self-report instruments (Kaye, Malone, & Wall, 2017) to measure for instance user and consumer experience (e.g. www.emojiscore.com). Since facial expressions can communicate a wide variety of both basic and complex emotions emoji-based selfreport tools may also afford the measurement and expression of mixed (complex) emotions that are otherwise hard to verbalize (Elder, 2018). However, while facial images and emoji are processed in a largely equivalent manner, suggesting that some non-verbal aspects of emoji are processed automatically, further research is required to establish whether they are also emotionally appraised on an implicit level (Kaye et al., 2021)."

2. How numerical values were determined

Thank you for pointing out this omission. We now include the following explanation of the scaling in the section on data analysis:

"The response data (i.e., the horizontal or valence and vertical or arousal coordinates of the check marks on the EmojiGrid) were quantified as integers between 0 and 550 (the size of the square EmojiGrid in pixels), and then scaled between 1 and 9 for comparison with the results of Yang et al. (2018) obtained with a 9-point SAM scale (Experiment I), or between 0 and 8 for comparison with the results of Aguado et al. (2018), also obtained with a 9-point SAM scale (Experiment II)."

3. Contribution and limitations

We now address some limitations of the present study (e.g. related to the measurement of

mixed emotions and the study design itself) in the Discussion section (see also our reply to the comments of dr Pham).

4. Minor points We replaced "persons" by "participants" throughout the text.

Competing Interests: The authors declare no competing interests.

Reviewer Report 01 September 2020

https://doi.org/10.5256/f1000research.27685.r69208

© **2020 Phan W.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

? Wei Ming Jonathan Phan

California State University, Long Beach, Long Beach, CA, USA

Thank you for the opportunity to review the manuscript: "Affective rating of audio and video clips using the EmojiGrid." This paper is primarily focused on validating the extension of a scale format (EmojiGrid) to a broader range of stimuli (audio and video). Overall, the paper makes some useful methodological contributions such as (1) the potentially greater ease for respondents for rating their emotions; (2) capturing both arousal and valence simultaneously; and (3) the use of more familiar contemporary symbols (emojis) compared to the SAM (Bradley & Lang, 1994).¹ I do have a few suggestions and concerns regarding the paper.

1. Limitation of the EmojiGrid in measuring single discrete emotions.

The EmojiGrid is useful for respondents when selecting which area of the grid corresponds to their current felt emotion. However, emotions are not bipolar in nature and can often co-occur together, e.g., feeling bitter-sweet (Larsen *et al.*, 2001; Larsen & McGraw, 2014)^{2,3}. Thus, the current form of the EmojiGrid is limited to assessing stimuli that invoke single discrete emotions and may not be as suited for assessing real-time affective reactions (e.g., to entertainment or news). This limitation can potentially be highlighted in the discussion. Importantly, this limitation can be solved by future and different operationalizations of the grid structure when mixed emotions are the object of inquiry.

2. Details regarding the stimuli selected.

Related to the first point, I note that the majority of the stimuli in both experiments (in particular experiment 1) seem to have a moderate amount of valence and arousal. Without knowing which stimuli were used, it is difficult to assess whether the emotion felt by the respondent was truly neutral or potential mix of emotions. To help the reader, please include two things potentially using tables in the supplementary material if needed. First, a greater description of which stimuli selected was expected to invoke which emotion in terms of both valence and arousal for both experiments. Second, please use a different numbering/labeling/coloring scheme that

corresponds to the stimuli instead of dots for figures 1 and 2 when comparing the results from this study to previous work. Both are important because it allows the reader to visually assess the extent an expected emotion of stimuli (e.g., high arousal and positive valence) truly maps onto the mean scores and for the potential discrepancy between the two scale formats for the same stimuli to be obvious. This is important for replication but also because there is a greater dispersion when the SAM rating format is used.

3. Comparing current data and alternate (future) research design.

When comparing data from the current experiments to previous experiments the regression estimates are locally optimized based on the sample used to generate them. Thus, a caveat and clarification to potentially include are that the comparisons made are akin to that of two independent samples. Relatedly, an alternate design to consider would be doing a 4-block repeated measures design. Where participants rate the same stimuli using the two rating formats twice as:

- 1. A then A
- 2. B then B
- 3. A then B
- 4. B then A

Blocks 3 and 4 would allow more direct comparisons between two different rating formats, especially given the greater dispersion in ratings observed when the SAM format is used.

4. Free response clicks within the EmojiGrid

I note that participants are free to click anywhere within the space of the EmojiGrid. I am curious as to variability/freedom that having no fixed anchor points generates. When participants respond do they more typically engage in: (1) subconsciously select a point close to one of the 25 potential points implied by the 5 X 5 grid of emojis, or (2) freely select a space with the grid, e.g., selecting point that corresponds to 2.30 arousal and 5.80 in valence? I ask this because the reliability of a scale is linked to the number of response points available (Preston & Colman, 2000; Schutz & Rucker, 1975)^{4,5}. If respondents are truly giving their ratings as (2) then greater reliability would be a potential additional advantage of using the EmojiGrid. If it were (1) the design of the EmojiGrid could include finer lines (i.e., more grid lines) to help respondents more easily locate their emotions on the Grid.

Minor points

- 1. Nationality information was collected from participants how was this information used? What was the distribution of nationalities for the participants?
- 2. I appreciate the way the authors determined their sample sizes.

I enjoyed reading your paper and hope you will find my comments helpful!

References

1. Bradley M, Lang P: Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*. 1994; **25** (1): 49-59 Publisher Full Text

2. Larsen J, McGraw A, Cacioppo J: Can people feel happy and sad at the same time?. *Journal of Personality and Social Psychology*. 2001; **81** (4): 684-696 Publisher Full Text

3. Larsen J, McGraw A: The Case for Mixed Emotions. Social and Personality Psychology Compass.

2014; 8 (6): 263-274 Publisher Full Text

4. Preston C, Colman A: Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*. 2000; **104** (1): 1-15 Publisher Full Text

5. Schutz H, Rucker M: A Comparison of Variable Configurations Across Scale Lengths: An Empirical Study'. *Educational and Psychological Measurement*. 1975; **35** (2): 319-324 Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature? $\ensuremath{\mathsf{Yes}}$

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others? Partly

If applicable, is the statistical analysis and its interpretation appropriate? Partly

Are all the source data underlying the results available to ensure full reproducibility? $\ensuremath{\mathsf{Yes}}$

Are the conclusions drawn adequately supported by the results? Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Survey methodology, rating formats, Emotions, and Emojis.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 10 Mar 2021 Alexander Toet, TNO, Soesterberg, The Netherlands

Dear Dr Phan,

Thank you for your helpful suggestions and constructive remarks, which helped us to improve the quality of our initial draft paper. In addition, we appreciate the fact that you spent your valuable time on this review.

1. Mixed emotions

We thank the reviewer for raising this important issue. We now address this limitation in the Conclusion section as follows:

"A limitation of the EmojiGrid is the fact that it is based on the circumplex model of affect which posits that positive and negative feelings are mutually exclusive (Russell, 1980). Hence, in its present form, and similar to other affective self-report tools like the SAM or VAS scales, the EmojiGrid only allows the measurement of a single emotion at a time. However, emotions are not strictly bipolar and two or more same or opposite valenced emotions can co-occur together (Larsen & McGraw, 2014; Larsen, McGraw, & Cacioppo, 2001). Mixed emotions consisting of opposite feelings can in principle be registered with the EmojiGrid by allowing participants to enter multiple responses. "

2. Stimulus details

Thank you for bringing this limitation to our attention. We agree that a labelling scheme (e.g. using the original stimuli identifier) makes a visual comparison between the experiments much easier. We therefore replaced the original graph with labelled graphs in the paper to allows the readers to visually assess and verify the expected emotions induced by the stimuli. We also added correlation plots for the mean valence and arousal ratings obtained both with the SAM and EmojiGrid to enable a direct comparison within each dimension.

In addition, we now also provide a more detailed description of the selected stimuli in a new set of Excel notebooks that we uploaded to the Open Science Framework These notebooks include a brief description of the nature and content of all stimuli, their original affective classification, and their mean valence and arousal values (1) as provided by the authors of the (sound and video) databases and (2) as measured in this study.

The notebooks also contain several graphs in which each of the stimuli is represented by the index number for easy identification. The graphs include plots showing (1) the relation between the mean valence measures obtained with the SAM and EmojiGrid, (2) the relation between the mean arousal measures obtained with the SAM and EmojiGrid, (3) the relation between the mean valence and arousal measures obtained with the SAM, and (4) the relation between the mean valence and arousal measures obtained with the EmojiGrid.

3. Comparing data

Thank you for pointing out this limitation. We now address this issue in the Conclusion section as follows:

"Another limitation of this study is the fact that the comparison of the SAM and EmojiGrid ratings were based on ratings from different populations (akin to a comparison of two independent samples). Hence, our current regression estimates are optimized based on the particular samples that were used. Future studies should investigate a design in which the same participants use both self-report tools to rate the same set of stimuli. "

4. Free response clicks

Thank you for raising this potentially important issue. We plotted the raw response data for visual inspection. The overall response pattern appears truly random and shows no regularities or evidence for attraction to any of the individual emojis lining the grid area or to any of the grid lines inside the grid area.

Minor points

Thank you for drawing our attention to this omission. All participants had the UK nationality. We now report this in the text. Nationality information was collected to check if the participants adhered to the recruitment restrictions as specified through Prolific. Thank you for your positive appraisal. Your comments were quite valuable, and definitely served to improve the quality of our paper.

Competing Interests: The authors declare no competing interests.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000 Research