


Please cite the Published Version

Diconne, Kathrin, Kountouriotis, Georgios K, Paltoglou, Aspasia E, Parker, Andrew  and Hostler, Thomas J (2022) Presenting KAPODI – The Searchable Database of Emotional Stimuli Sets. *Emotion Review*, 14 (1). pp. 84-95. ISSN 1754-0739

DOI: <https://doi.org/10.1177/17540739211072803>

Publisher: SAGE Publications

Version: Published Version

Downloaded from: <https://e-space.mmu.ac.uk/629122/>

Usage rights:  [Creative Commons: Attribution 4.0](https://creativecommons.org/licenses/by/4.0/)

Additional Information: This is an Open Access article published in *Emotion Review* by Sage.

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)



Presenting KAPODI – The Searchable Database of Emotional Stimuli Sets

Kathrin Diconne , Georgios K. Kountouriotis, Aspasia E. Paltoglou, Andrew Parker and Thomas J. Hostler 

Department of Psychology, Manchester Metropolitan University

Abstract

Emotional stimuli such as images, words, or video clips are often used in studies researching emotion. New sets are continuously being published, creating an immense number of available sets and complicating the task for researchers who are looking for suitable stimuli. This paper presents the KAPODI-database of emotional stimuli sets that are freely available or available upon request. Over 45 aspects including over 25 key set characteristics have been extracted and listed for each set. The database facilitates finding of and comparison between individual sets. It currently contains sets published between 1963 and 2020. A searchable online version (<https://airtable.com/shrnVoUZrwu6riP9b>) allows users to select specific set characteristics and to find matching sets accordingly, as well as to add new published sets.

Keywords

emotion, stimuli sets, database, systematic review

Date received: 29 January 2021; revised: 19 October 2021; accepted: 19 December 2021

Introduction

One of the most challenging aspects of designing a study involving emotions can be the selection of suitable emotional stimuli (ES) for the research goals. The aim of the present paper is to reduce this challenge and assist the researcher in making a more informed choice about which stimulus set best suits such goals, via the creation of a new tool – the searchable KATHrin POs. DIconne (KAPODI) database of emotional stimuli.

Choosing appropriate stimuli is no simple matter as the range of examples of ES include images, words, music, speech, or even video-clips. Their application is diverse; for example, word- and speech ES have been used to aid diagnoses in a clinical context (Haro et al., 2017; Nieuwenhuis-Mark et al., 2009), while video and image stimuli have been applied in therapy for alcohol addiction (Pronk et al., 2015), and eating disorders/obesity (Miccoli et al., 2014, 2016). Furthermore, emotional speech and music stimuli have been used for advertisements (Zander, 2006), learning aids (Schön et al., 2008), and political and

linguistic research (Cullen & Harte, 2018; Edelman et al., 1992). Images of facial expressions are also useful for training emotion recognition and emotion expression in children (Cardos et al., 2017), individuals with autism (Wingenbach et al., 2016), and schizophrenia (Gutiérrez-Maldonado et al., 2014). Additionally, ES can be used in intercultural research (Sacco et al., 2016), human-machine interaction (Battocchi et al., 2005), and machine learning (Zafeiriou et al., 2016) within artificial intelligence. They are also useful for automatic recognition of anger in speech (Neiberg & Elenius, 2008), for instance, to indicate the customers' emotion to staff in call-centres, as well as for emotion detection within texts such as online messages, comments on websites, and blog entries (e.g., Ramalingam et al., 2018). Finally, they can be used for automatic emotion recognition in human movement (Crane & Gross, 2007) such as detecting aggression, which can be applied within video surveillance of prisons or public spaces for security reasons. The ability to detect emotions automatically is thus of great interest to many organizations, including private businesses and

the government. For algorithms training machines, ES are used as the ground truth.

Therefore, as well as the actual ES themselves, validation data is just as important. These assessed rating data provided along with stimuli are called *normative* data. This type of information allows the researcher to make reasoned decisions about which ES set to use. However, all too often, comparison between stimuli sets is impeded by variations in how emotion is characterized and validated. More specifically, a variety of assessment approaches have been used to validate the items including the type of rating scale, its length, and the characteristics of the population upon which the stimuli were normed. For instance, researchers may use a *dimensional* approach (e.g., valence, arousal, dominance; Osgood, 1952; Osgood et al., 1957) or a *categorical* approach (e.g., *the big six*, namely happiness, sadness, anger, fear, disgust, and surprise; [Ekman et al., 1969]); they may apply the Self-Assessment Manikin (SAM) scale (Bradley & Lang, 1994; Lang, 1980), the Likert-scale (Likert, 1932), or the visual analogue scale (Hayes & Patterson, 1921). As these scales often vary regarding length and assessed characteristics (e.g., distinct emotion such as happiness or sadness, dimension such as valence or arousal; Ferré et al., 2012; Imbir, 2015, 2016; Katsimerou et al., 2016; Miccoli et al., 2016; Provost et al., 2015), it is important to establish a means of comparison across sets to provide a basis for a more effective selection for the research objectives.

Commensurate with a rapidly growing interest in emotion research (e.g., emotion regulation [Gross, 2015] and emotional development [Pollak et al., 2019]), the need for ES has also increased. This is reflected in the extensive number of ES sets that have been steadily developed over the past few decades which have increased and broadened the range of available stimuli. The result is that when attempting to find suitable ES, one is faced with the daunting task of evaluating a vast number of sets which can be time consuming and unrewarding. Additionally, some sets may be outdated, difficult to access, or unavailable.

Interestingly, despite the abundance of ES, emotion-related studies tend to use only a few well-known stimuli sets that have, as a result, established research credentials. Examples of these include the *Affective Norms for English Words* (ANEW; Bradley & Lang, 1999), the *International Affective Picture System* (IAPS; Lang et al., 1997), the *NimStim Set of Facial Expressions* (NimStim; Tottenham et al., 2009), and the film set by Gross and Levenson (1995). Indeed, their renown may overshadow smaller, recently published, or lesser-known sets, and certain stimuli may have become overly familiar to subjects. Therefore, new fully normed stimuli sets that are less familiar to participants are constantly required and being developed.

Nevertheless, relying on normative data can be a double-edged sword. For instance, although researchers may assume that selected stimuli will possess the characteristics specified by the normative data to bring about the intended effects, the

ratings themselves may be affected by a plethora of external factors. These could for instance be the participants' age (Grandy et al., 2020; Isaacowitz et al., 2007), gender (Hall & Matsumoto, 2004; Lithari et al., 2010; Nater et al., 2006), ethnicity (DeBusk & Austin, 2011; Deng et al., 2017), mental health (Mathews & MacLeod, 2005), hormone levels (Little, 2013), or social and cultural background and environment (Boiger et al., 2018; Kring & Gordon, 1998; Matsumoto et al., 2008). Furthermore, stimuli characteristics such as stimuli model's gender (Adolph & Alpers, 2010), or study characteristics such as order of stimuli presentation (Thayer, 1980a, 1980b), and evaluation context (Delatorre et al., 2019) can also influence the perception of ES. In other words, a stimulus can elicit different emotions or physiological reactions between different individuals, or even within one individual, and is also dependent on the context (Frijda, 2007).

Hence, a set validated for one population or context may not have the same effect - and thus not be valid - in another population or context. For example, research investigating the emotional perception of images in countries suffering from violence showed that Israeli adults rated images differently than adults in the United States (Okon-Singer et al., 2011). Moreover, the year the study was conducted may play an important role: For example, an image of the World Trade Center in New York City, presented to participants prior to the 11 September 2001, may have elicited different emotions compared to after that date. Images or video stimuli may include cues such as hair style or fashion that can easily be associated with a specific decade and thus seem outdated when seen today.

Unfortunately, not all available stimuli sets document these factors. Therefore, it is important to consider small details of ES set construction such as characteristics of the rating population, date of created stimuli, or country of research, when selecting stimuli and/or relying on normative data. Hence, the availability of a large pool and broad range of assessed ES is of great importance.

Although short overviews of existing sets have previously been provided in the research literature, these are typically not comprehensive (e.g., Grünh & Sharifian, 2016; Krumhuber et al., 2017), as they have not systematically searched for sets, or the overviews focus only on specific types of ES such as *words* (e.g., Riegel et al., 2015; Scott et al., 2019) or *faces* (e.g., Prada et al., 2018; Tu et al., 2018).

Therefore, a step forward in establishing a more extensive and inclusive tool for emotion researchers is a central searchable database that can provide access to pre-validated stimuli and that lists key characteristics such as gender, age, and ethnicity of the rating population, type and length of rating scales, and assessed emotions and dimensions. It will enable researchers to search for suitable stimuli, facilitate comparison across sets, and save time. Moreover, such a database would be useful in allowing more effective study replication as well as more successful manipulation or

control of a multitude of external (e.g., country of the survey, year of the study, ethnic background of participants) as well as internal (e.g., luminance, color, duration, or video/audio speed of presented stimuli) factors that are important in emotion research.

With no existing comprehensive review to date, the objective of this study was to systematically review existing and freely available sets of ES and provide an overview by

documenting the central characteristics of each set. A searchable online version of the results will serve as a tool allowing specific set characteristics to be selected and display filtered results. The hope is that the resulting database will be a useful resource for researchers planning studies and in need of stimuli and/or assessed emotional rating data. Additionally, it may be directly beneficial to other contexts such as in therapeutic settings, or human-machine interaction.

Table 1. Coded characteristics for each subfolder.

Coded characteristics	Subfolder					
	Audio	Faces	Images	Video	Words	Mixed
Title of publication	yes	yes	yes	yes	yes	yes
Authors	yes	yes	yes	yes	yes	yes
Year of publication	yes	yes	yes	yes	yes	yes
APA citation	yes	yes	yes	yes	yes	yes
University affiliation	yes	yes	yes	yes	yes	yes
Stimuli set name	yes	yes	yes	yes	yes	yes
Type of stimuli	yes	yes	yes	yes	yes	yes
Resolution	yes	yes	yes	yes	-	yes
Content	yes	yes	yes	yes	yes	yes
Expression authenticity	yes	yes	-	-	-	yes
Ethnicity	-	yes	yes	yes	-	yes
Number of stimuli	yes	yes	yes	yes	yes	yes
Stimuli length	yes	yes	-	yes	yes	yes
Number of models	yes	yes	-	yes	-	yes
Sex of models	yes	yes	-	-	-	yes
Age of models	yes	yes	-	-	-	yes
Specific number of emotions	yes	yes	yes	yes	yes	yes
Color / hue	-	yes	yes	yes	-	yes
Language	yes	-	-	yes	yes	yes
Categorical approach	yes	yes	yes	yes	yes	yes
Dimensional approach	yes	yes	yes	yes	yes	yes
Categorisation	yes	yes	yes	yes	yes	yes
Rating scale(s)	yes	yes	yes	yes	yes	yes
Additional Information	yes	yes	yes	yes	yes	yes
Included specific emotions	yes	yes	yes	yes	yes	yes
Number of included specific emotions	yes	yes	yes	yes	yes	yes
Validation for subgroup	yes	yes	yes	yes	yes	yes
Rating by	yes	yes	yes	yes	yes	yes
Student/non-student raters	yes	yes	yes	yes	yes	yes
Country of study	yes	yes	yes	yes	yes	yes
Continent of study	yes	yes	yes	yes	yes	yes
Number of raters	yes	yes	yes	yes	yes	yes
Number of raters per stimulus	yes	yes	yes	yes	yes	yes
Source access	yes	yes	yes	yes	yes	yes

Note. University affiliation = author-affiliation as indicated in the publication; content = short description of the database content; expression authenticity = e.g., natural vs. acted or posed, applicable to only emotional expressions (i.e. vocalisations or faces); categorical approach = stimuli assessed after the categorical approach of emotion; dimensional approach = stimuli assessed after the dimensional approach of emotion; categorisation = assessed dimensions; additional information = information that seems important that could not be coded otherwise; validation for subgroup = sets that have specifically been validated for a subgroup/ information for certain research fields are proposed by authors; rating by = characteristics of stimuli raters; student/non-student raters = coded whether raters were university/college students.

Method

The methodological procedure of the current systematic literature review consisted of three main stages: first, conducting a systematic literature review aiming to detect all existing ES sets; second, coding and listing of key characteristics of all included sets; and third, creating a searchable online database.

Stage One: Systematic Literature Review

To capture the greatest possible number of papers proposing ES sets, an appropriate keyword selection had to be made and inclusion as well as exclusion criteria determined. In two consecutive steps, publications not meeting inclusion criteria were excluded.

Information sources. The keyword search was conducted in April 2019, on *PsychInfo*, *Medline* (EBSCOhost), and *Web of Science*. The time frame for publication date was set to 1950 to 2019 for PsychInfo and Medline, as well as 1970 to 2019 for Web of Science, respectively, as the early beginnings of emotion research and proposal of stimuli sets can be pinpointed approximately to the 1950s (Osgood, 1952). The reduced time frame for Web of Science was restricted by the database entry options.

Due to a long time-spanning coding process after the first search, the same search was conducted a second time in June 2020 aiming to detect all studies presenting ES published between January 2019 and June 2020. Literature search included emotion- and stimuli-related keywords and were kept limited to the six basic emotions (Ekman et al., 1969). The applied exact keywords were: “(emotional) OR (emotion) OR (affect) OR (affective) OR (fear) OR (disgust) OR (happiness) OR (anger) OR (angry) OR (sad) OR (sadness) OR (surprise) [IN all text] AND (stimulus OR stimuli OR picture\$ OR word\$ OR video\$ OR audio OR film\$ OR sentence\$) [IN all text] AND (set OR database OR list OR library OR norms) [IN title]”.

The keyword *database* could only be searched for [IN title], as many search-engine databases commonly include this term below a paper’s abstract, leading to tens of thousands keyword search results when searched for [IN all text] or [IN abstract].

Eligibility criteria. Papers were selected according to the criteria below:

To be included, papers had to (I) be peer-reviewed; (II) be published in English, French, or German; (III) be published between 1950 and 2020; (IV) include ES that are either (a) images; (b) video; (c) audio; or (d) words; and (V) include sets freely accessible to the research community.

Excluded were all sets containing ES such as heat, pressure, or odor, as well as ES created for animal studies. Further, sets providing solely emotional physiological data unaccompanied by stimuli were excluded.

As discussed earlier, various factors can influence the perception of ES; therefore, in this systematic literature review validation of presented stimuli through participants was not considered a prerequisite. Publication without validation was sometimes the case for studies where models were asked to express certain emotions, (e.g., Minear & Park, 2004; O'Toole et al., 2005; Yingliang et al., 2006), or also for word lists created by the researchers themselves (Barrington, 1963).

Study selection. Literature search results were uploaded to Rayyan Software (Ouzzani et al., 2016), an internet-based software program that facilitates the study selection process. All search results were manually and independently screened against inclusion criteria, based on title and abstract. Uncertainties concerning inclusion were resolved through discussion and consensus of two to four researchers. When necessary, additional information was sought directly from study authors. Concerning inquiries regarding availability, authors were contacted twice via e-mail within approximately two to three months. If they did not respond after the second enquiry, the set was considered unavailable.

Stage Two: Data Collection Process and Data Items

Each publication that met inclusion criteria was then read two times independently and assigned to one of six subfolders: (1) *audio*, (2) *faces*, (3) *images*, (4) *video*, (5) *words*, and (6) *mixed*, depending on the type of most stimuli included in the presented set. Characteristics (e.g., year of the publication, type of stimuli, number of stimuli, resolution, number of raters, or applied rating scales) were coded. A detailed outline of these characteristics can be found in Table 1. Whenever information was not provided in the paper, this was noted as not available (n/a). In case of information inconsistencies within the publication, or to resolve any uncertainties, study authors were contacted.

Stage Three: Creation of a Searchable Online Database

All extracted information was coded in an Excel sheet. An online version of the database was created. It serves as a search tool in which specific criteria such as type of stimuli, models' age, rating scales, included emotions, etc.

can be selected, leading to the display of solely sets containing these characteristics. The searchable database can be found at <https://airtable.com/shrnVoUZrWu6riP9b>.

Results

The first keyword search yielded 1,877 pieces of published work (443 in PsychInfo, 393 in Medline, and 1,041 in Web of Science). Duplicates ($n=616$) were removed, 1,261 search results remained for manual scanning. This manual scanning was conducted in two subsequent steps: first, a coarse selection based on title and abstract, then thorough reading of the full publication. Based on title and abstract, $n=951$ results were excluded due to unrelated content (e.g., studies on animals, publications originating from chemistry or physics), leaving 310 publications for thorough reading. In this second step, another 73 publications were excluded because (a) their content was not relevant to the systematic review ($n=56$), or (b) the described set was not available to the research community/authors did not respond to e-mail requests concerning availability of the set ($n=17$). A more detailed overview of the individual steps can be found in the PRISMA flow diagram (Moher et al., 2009) in Figure 1.

A few publications mentioned further stimuli sets that were not detected by our initial key word search. Therefore, we decided to further verify the availability of all sets presented in tables within the publications that were already part of our database (extended table-search). All additional sets conforming to the inclusion criteria as mentioned above were included: Another $n=74$ publications were added. At this point, a total of 311 publications had been included and coded. The same search and selection process were conducted for the second updated search covering all studies published between January 2019 and June 2020.

This updated keyword search yielded another $N=268$ results (34 in PsychInfo, 162 in Medline, and 72 in Web of Science). Duplicates within these three search databases ($n=78$), as well as duplicates of results from the first search ($n=36$) were removed; 154 papers remained for manual scanning. A further $n=106$ publications were removed based on title and abstract; a further $n=15$ papers were removed after thorough reading. One paper was excluded, as authors did not respond to the request regarding set availability. A total of $N=53$ publications were added through this second updated search: an initial $n=32$ publications, and an additional $n=21$ publications from the extended table-search.

With 311 publications from the first search and 53 publications from the updated search, at the point of creation the database contains a total of $N=364$ publications. Each publication presents at least one set of ES and/or new assessed rating data. All publications and their extracted main criteria are listed in an Excel spreadsheet available as *Supplementary Material*, and an online version of the database is also

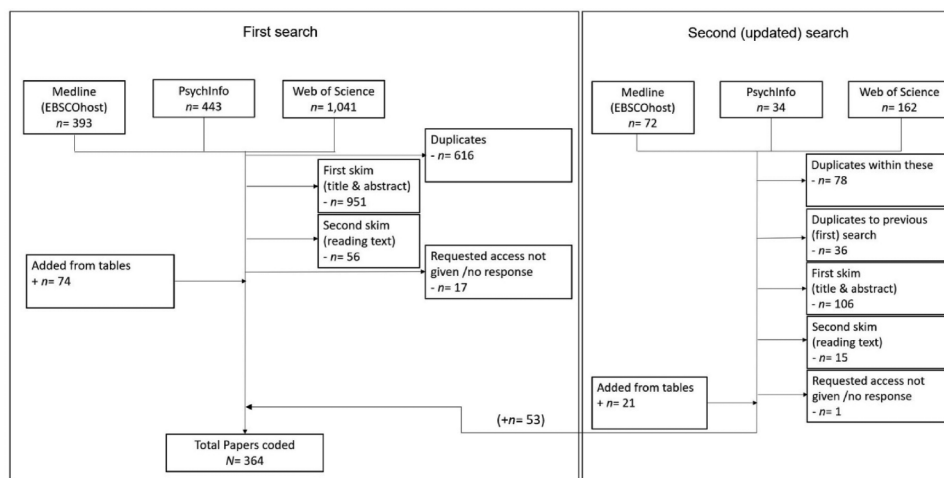


Figure 1. PRISMA flow diagram of the research procedure. *Note.* Research procedure conducted applying key words in three databases (PsychInfo, Web of Science, and Medline (EBSCOhost)). First search conducted in April 2019, and the second (updated) search conducted in June 2020.

available. Note that the supplementary material contains information only up to 2020, while the online version of the database will keep being updated.

In the following section, a brief explanation of the subfolders is included. Currently, the database contains publications from 1963 to 2020. Note that the final keyword search in this study was conducted in June 2020, therefore stimuli sets published after that date are not included in the discussion below.

Stimuli Subfolders

For easier comparison across stimuli, each publication was allocated to one of six subfolders describing the type of stimuli: (1) audio ($N=35$), (2) faces ($N=117$), (3) images ($N=35$), (4) video ($N=43$), (5) words ($N=89$), and (6) mixed ($N=45$). Note, that the number of *publications* is named, as some publications present more than one set (e.g., *FEEDBver.1* & *FEEDB*, Szwoch, 2014; *ATAL* & *ETAL*, Torkamani-Azar et al., 2019). Publications were allocated to the best fitting sub-folder, which means that the included stimuli types are not exclusive: For example, *face* stimuli, although mostly presented as images, were allocated to a separate folder and not the images folder. It is also important to mention that some sets (partially or in their entirety) have been translated into another language, validated in another country, or validated for a different age-group, and thus do not contain *new* stimuli. However, they present new normative rating data. Such examples are the translation of the ANEW set (Bradley & Lang, 1999) into Italian (Montefinese et al., 2014) and Portuguese (Soares et al., 2012); the validation of the IAPS (Lang et al., 1997) for a Brazilian population (Ribeiro et al., 2005), a population of countries suffering from violence (Okon-Singer et al., 2011), or the creation of a subset relevant for Borderline

Personality Disorder (Sloan et al., 2010); the creation of an audio version of an existing word set for cross-modal validation (Kanske & Kotz, 2011) or clinical subsamples (Kanske & Kotz, 2012). Examples of the validation of stimuli for different age-groups can be found with the *Besançon Affective Picture Set* (BAPS) (*BAPS-Ado*, Szymanska et al., 2015; *BAPS-Adult*, Szymanska et al., 2019) and other duplicated sets include those with modification of stimuli, for example by morphing existing face-images and creating dynamic stimuli, as can be found with the *FACES* set (Ebner et al., 2010) modified into *DynamicFACES* (Holland et al., 2019), and the *KDEF* set (Bartlett et al., 1999) modified into *KDEF-dyn* (Calvo et al., 2018).

A short description of results per subfolder follows:

Audio. Audio-stimuli contain spoken individual words, sentences, pseudo language/gibberish, as well as music. The focus of emotion varies between intonation and semantic content of stimuli. This means that stimuli can be selected with a focus on perception of emotional tone, or emotional content, or a combination of both for example by using stimuli with emotional semantic content expressed in various emotional tones.

Faces. A specific focus within video and image stimuli are facial expressions. This subfolder contains the most publications in relation to the other folders and accounts for almost one third (32.14%) of all sets. A growing interest in automatic emotion detection within artificial intelligence and progress in human-machine interaction are pushing researchers to continuously adjust and improve algorithms; these strongly depend on standardized and/or validated sets of face-stimuli. Proposed sets contain both grey-scale stimuli and color-stimuli, in 2D as well as 3D. Further, still-

image stimuli, as well as video and dynamic stimuli are included. Dynamic stimuli can be constructed artificially based on individual images. Similarly, this folder also includes sets with morphed stimuli which are stimuli that are created by superimposing multiple images or video frames. In some cases, researchers provide video streams, as well as individual frames (image-stimuli) of the recorded videos. Standardization such as fixed pupil position across all models from one set, or removal of potentially distracting information (e.g., jewelry, hair, makeup, clothes) is coded. Both posed, as well as natural/spontaneous emotion expressions are included. Sets also vary in the degree of homogeneity regarding models' gender, age, and ethnicity.

Images. Image-stimuli in this database cover a variety of subjects. Examples are beverage and food images, natural scenes, threat and crime stimuli, as well as line drawings of social contexts. Image sets were often created for research in specific subgroups such as alcohol addiction, eating disorders, or phobia.

Video. Video-stimuli are one of the most effective type of stimuli for emotion elicitation (Gross & Levenson, 1995; Westermann et al., 1996) and thus are often used in studies aiming to induce a specific emotion in participants. Stimuli vary from video-only, to audio-visual stimuli that are accompanied by speech or music. Included are sets that present video clips extracted from movies and TV-shows, clips recorded specifically for the study, as well as motion-capture data where only point-lights representing body-part position in space are visible. Especially when stimuli are extracted from movies or TV-shows, quality and camera angle as well as microphone sources can vary widely within one set.

Words. Word-stimuli sets are stimuli that range from written individual words, to sentences or text passages. Similar to audio-stimuli, sets include real words in various languages, as well as gibberish speech. Multiple studies report extracting a selection or using all words from an already existing set such as the ANEW (e.g., Nieuwenhuis-Mark et al., 2009; Schmidtke et al., 2014; Sutton & Altarriba, 2016).

Mixed. The *mixed*-subfolder was created to list all stimuli sets that cannot clearly be allocated to any of the other subfolders. Often, these sets contain a combination of different types of stimuli or additionally provide physiological data such as respiration, heart rate, skin conductance, or temperature. Physiological data recordings can be extremely valuable to research focussing on emotion recognition through artificial intelligence and to the understanding of physiological processes during the experience of different emotions. This sort of data, for instance, can be useful for the investigation of emotion regulation or understanding of disabilities in relation to emotion such as anxiety, apathy, or psychopathology.

Navigation of the Database

In addition to listing all emotional stimuli sets with their key characteristics as presented in the Excel sheet version (see *Supplementary Material*), the online version of the KAPODI database includes a search tool allowing the selection of stimuli sets according to specific criteria. Based on the stimuli type (e.g., audio, faces, images, video, words, mixed), each set is listed in one of the six subfolders (see Figure 2, [a]). Further, the selection between gallery or grid view presents the stimuli sets in a list (grid view; Figure 2, [b]), or as individual cards that can be selected for more detailed information (gallery view; Figure 3, [b]).

Within each subfolder (e.g., audio stimuli), (see Figure 3, [a]) the filter tab [c] allows the selection of stimuli sets based on the specific filter(s) (e.g., language, [d]) and a refined search (e.g., English and Japanese, [e]). All extracted key characteristics mentioned in Table 1 can be set as a filter. Moreover, users may search sets by entering key words in the search bar. Only sets matching the search criteria are displayed to the viewer.

Discussion

The searchable KAPODI database was constructed in response to an increasing number of publications providing ES sets to the research community. The numerous publications reflect a growing demand for tailored stimuli as well as a growing interest in the field of emotion research. In the present work, the first comprehensive systematic review of emotional stimuli sets was conducted, and key set characteristics were coded to allow for comparison between sets and aid researchers in choosing appropriate sets for their research. The resulting KAPODI database contains the largest list of available stimuli sets to this date and is therefore a useful contribution to research on emotion and beyond.

In the following section, the database will be discussed regarding its use and its limitations. Finally, the creation and publication of emotional stimuli sets will be discussed with recommendations regarding how these should be reported.

Using the Database

Researching the availability of different ES sets is a time-consuming task and it often leads researchers to resort to well-known and widely used stimuli sets despite a wide range of other stimuli sets being available. At the time of publication, the KAPODI database comprises 364 publications from 1963 to 2020 that cover various types of stimuli such as audio-, image-, video-, and word-stimuli, or a combination thereof (e.g., audio-visual). Six sub-folders were created within the database for easier comparison of similar stimuli and to facilitate stimuli search for researchers in the future. Over 25 key characteristics have been coded for

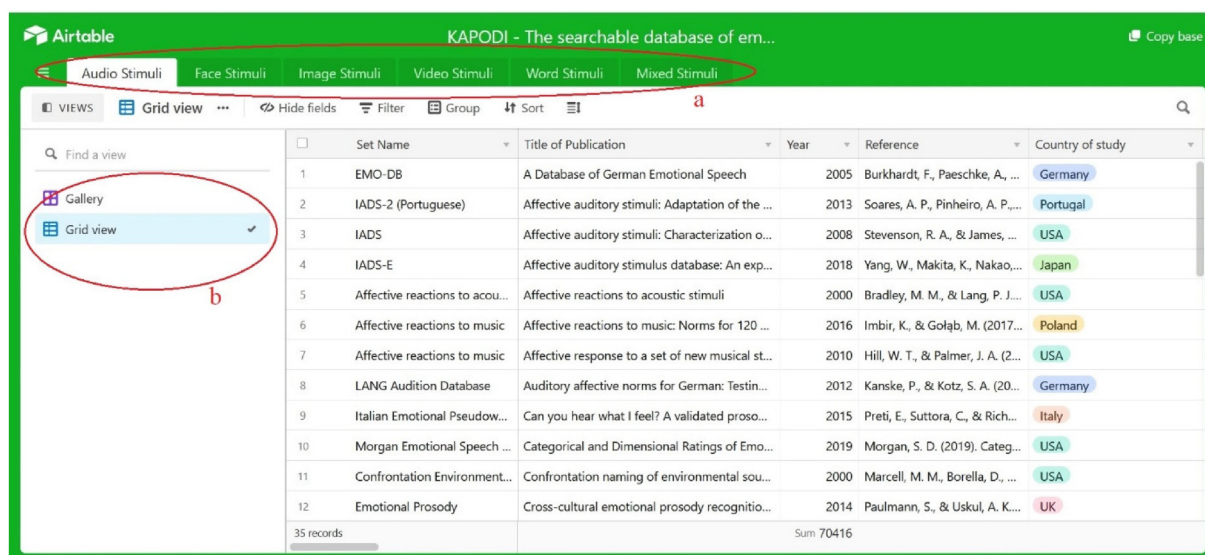


Figure 2. Exemplary view of the KAPODI searchable database I. *Note.* Stimuli sets are separated by type of stimuli (a) and can be viewed in a gallery or grid view (b).

each set, leading to over 45 criteria that researchers can use to find appropriate stimuli (e.g., the characteristic *rating scale* with the criteria SAM-scale, Likert scale, visual analogue scale, forced-choice, other). The database allows researchers to see whether a study has created a new set or has validated stimuli from a pre-existing set in a new population (e.g., different age group, different ethnicity, or different country). Set characteristics of interest can easily be compared, facilitating choice of set, and accelerating the research process.

Information coded about the included emotions may for instance allow researchers to select stimuli of a distinct emotion, or select stimuli rated as neutral, for comparison. Researchers may search for stimuli sets based on applied rating scales (e.g., SAM-scale, Likert scale, visual analogue scale) or length of used scales, which may be important information for choosing new stimuli for replication studies. Researchers can also select sets that include a minimum number of stimuli, a minimum number of included models (e.g., within face-stimuli), or choose stimuli sets that include a specific type of content (e.g., images of food, or fear-inducing images). Moreover, it is possible to search for stimuli of a specific language or sets that include an additional type of data (e.g., physiological recordings).

The coded number of raters per stimulus allows researchers to search for stimuli assessed by a minimum number of raters, which is relevant for discerning the reliability of the ratings. Additional information regarding the rating population allows selection of stimuli based on the type of assessors (e.g., sets normed with student populations or via crowdsourcing). For certain types of stimuli, further information is coded to allow researchers to find exactly what they are looking for. For example, a researcher may wish to find audio stimuli which include natural (non-acted) expressions

of happiness. The selected filters would be as follows: within the *audio-stimuli* subfolder, *expression authenticity is natural* (Filter 1), and *all included emotions, has any of, happiness* (Filter 2). The only dataset currently included in the database and meeting these search criteria is the OxVoc (Oxford Vocal Sounds Database; Parsons et al., 2014).

Finally, further information is provided regarding the context of creation for each ES (e.g., researchers' affiliation and country of study), as we believe that this may provide valuable information for researchers interested in the development of different types of stimuli from a geographical perspective.

In summary, the database provides the researcher with more flexibility in selecting an appropriate stimulus set and provides a systematic basis for going beyond classic ES sets (e.g., ANEW). This central database facilitates access to, and eases comparison between stimuli and/or sets for a wide range of applications and for researchers from a wide range of disciplines.

Strengths and Limitations

Despite its benefits for research, the database described is not without limitations. These, along with a few examples, are outlined in the following section.

Similar to the *Pictures of Facial Affect* presented by Ekman (POFA, Ekman, 1976), some researchers created stimuli for use in their study without consideration of the stimuli being used in further experiments by different researchers. Without doubt, each of the sets included in the KAPODI database has distinct strengths and was created for a specific aim, filling the gap in the availability of standardized stimuli. Nevertheless, depending on the initial research

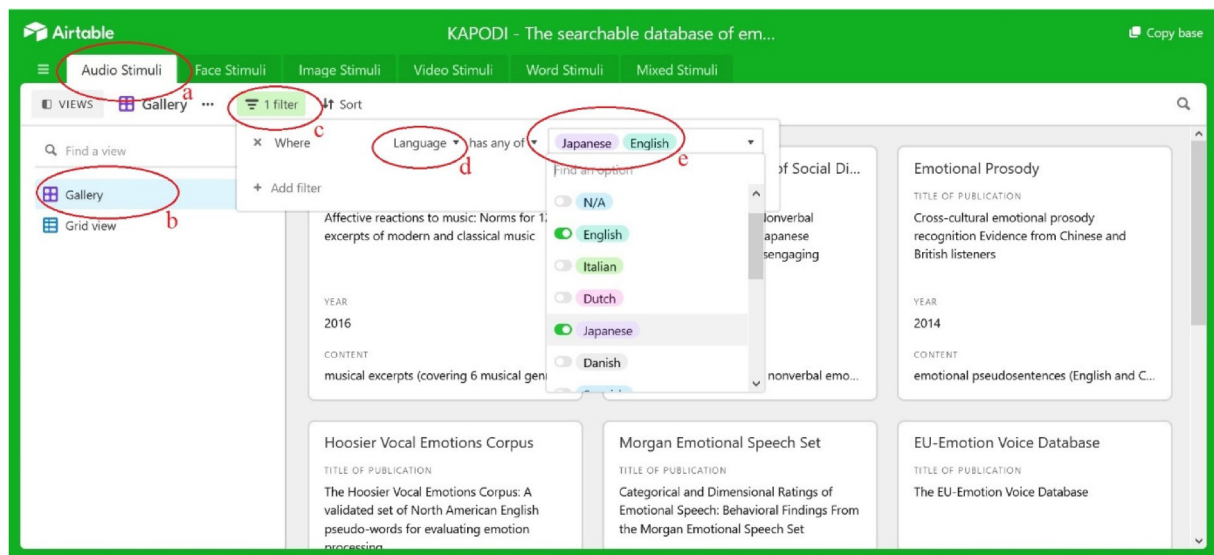


Figure 3. Exemplary view of the KAPODI searchable database II. *Note.* The audio subfolder is selected (a) and view set to gallery view (b); the filter (c) allows the selection of key set characteristics (d) (e.g., language) with a refined selection among all available options (e) of the set filter.

aim, different key characteristics were regarded important by the authors and therefore reported along with the stimuli, while others were not. This is a difficulty that is reflected through unclear or incomplete available information when comparing all included sets.

The root of this difficulty could be due to incompatible theories of emotion; for instance, there is no agreement for a unitary definition of emotion (Izard, 2007) to this date. Therefore, the current database employed categories that were thought to best represent the various facets of emotional information and allow comparison between sets without subscribing to a specific overall theory of emotion categorization. In summary, criteria from different sets were coded according to our best understanding of set content and study procedure, while also keeping in mind the usability of our created database (e.g., set search by key words) for interested researchers.

To give an example, some scientists accept only the *basic six* (Ekman et al., 1969) as emotions and reject any others. We, however, did not judge the different terms named *emotions* by other researchers, but rather listed them as proposed in the source article. As an illustration, in two studies, *smile* was mentioned as an expression (McDuff et al., 2019), or images have been classified/labelled according to smile (Samaria & Harter, 1994). These two studies were treated as exceptions and smile was listed as an emotion, so that these sets are also detected when searching through our database.

Comparably, other terms may have differed in some publications, though we did not modify terms while coding: despite most publications naming happiness as one of the basic emotions, in a few cases, happiness was replaced by *joy* (e.g., Costantini et al., 2014) or *amusement* (e.g., Yan

et al., 2013). Though this seemed to depend on translation from other languages (e.g., French or German: Bertels et al., 2014; and Hewig et al., 2005), it was necessary to find a consensus and create categories to facilitate the search within the database without changing the meaning of terms used in the original study. In this example regarding happiness/joy/amusement, the original authors' decision was accepted during the coding process. This means that in a few cases, joy may be listed as one of the basic emotions, while in other cases it was listed as an emotion that differs from happiness (e.g., Soleymani et al., 2012). The same applies to amusement.

Another example is that of the three dimensions *valence*, *arousal*, and *dominance*, on which stimuli are assessed in many publications. In a study conducted by Marcell et al. (2000), *pleasantness* was assessed. We coded this as the equivalent of valence. In another publication, *potency* was assessed (Kleinsmith et al., 2011) with authors mentioning that it is also referred to as *dominance*. However, given that the authors decided to use the term potency rather than dominance, we assumed that the term was chosen for a specific reason, and therefore we did not code it as dominance. Additionally, Schmidtke et al. (2014) suggest to differentiate between dominance and potency, as the latter 'mainly differs in its independence from the raters' perspective' (p. 1110).

Recommendations for the Creation and Publication of Stimuli Sets

To offer a large applicability of good quality ES, researchers creating and presenting stimuli in the future should generally consider three aspects: high-quality stimuli, good validation procedure, and clear reporting. That is, researchers should

1) aim to create high-quality stimuli (e.g., high resolution, or high number of frames per second for video recordings); 2) validate the created stimuli by a large sample size of a well-justified selection of assessors (e.g., assessors with the same main characteristics as potential target groups) which is especially important for stimuli created for specific target groups such as individuals with alcohol or food addiction; and 3) include and clearly communicate detailed technical information (e.g., color spectrum or luminance) regarding each stimulus.

The issue of clear reporting is particularly important, we will therefore expand on this. Especially among video stimuli, information regarding sex and age of models is frequently missing. Similarly, stimuli sets within the *mixed* folder frequently did not contain information regarding the included ethnicities in their video recordings. Furthermore, some sets did not include information regarding language or color of stimuli. Though individual sets may have been created for a specific research aim and therefore may have suited a specific survey design, missing information may limit the appropriate use of the stimulus set. Moreover, the absence of detailed information - especially regarding the stimuli and validation characteristics - complicate the interpretation of study effects.

More recently published stimuli sets are often good examples of comprehensive reporting and high-quality stimuli, reflecting an increasing understanding of the need for relevant information to be included, but also of improving technological abilities (e.g., *CAFE-set*, LoBue & Thrasher, 2015; *EU-Emotion Voice Database*, Lassalle et al., 2019; *Food-Cal*, Shankland et al., 2019).

A central aim regarding future research conducted in relation to ES sets should be to improve the uniformity in reporting the characteristics of the set. Hence, it is suggested that researchers developing stimuli sets in the future should include information regarding all the key characteristics established through this current systematic review. For guidance, researchers may use the KAPODI submission form (see the link in the *Final remarks* section) as a checklist when reporting information of their stimuli set. It is further suggested that established terminology is used (e.g., the dimensions valence, arousal, dominance), unless authors specifically wish to differentiate and justify their own language and terms.

Finally, researchers should ensure that their ES sets are made available freely and openly to other researchers, which will substantially contribute to transparency and reproducibility of research procedures (see Munafò et al., 2017).

With the central aim of supporting the efficiency of scientific research and knowledge accumulation, we took the decision to include solely stimuli sets that are publicly available/freely available upon request. The information regarding availability of the set was taken from the original source. Some publications include an internet link, directly leading to the freely accessible set or to a compliance form for researchers. Others provide an e-mail address through

which sets and data can be requested to the author(s) directly. Nevertheless, links and/or e-mail addresses may have changed, or sets may no longer be available. We therefore cannot guarantee that sets are truly available at present, even if stated so in the original source. Rather than relying on requests for access via email, we recommend that in the future, researchers should upload their created sets to a website or repository granting availability and easing access to stimuli to colleagues. An automatic validation system (for instance through a form requiring assurance regarding the academic purpose of accessing the set) could restrict access to researchers only. This further ensures access to the stimuli to remain the same, even if the set creator has changed or left their institution.

Final Remarks

Using a systematic review methodology, the current study aimed to identify as many available ES sets as possible. The resulting searchable database, which can be found on <https://airtable.com/shrnVoUZrwu6riP9b>, currently contains 364 different stimuli sets. It is available to the research community and all included stimuli sets are freely available or available upon request. By making all extracted and listed set key characteristics available in an Excel sheet as well as through the website [Airtable.com](https://airtable.com/shrnVoUZrwu6riP9b), we further ensure its permanent availability. We aim to maintain the searchable online version of the database updated. By offering researchers the opportunity to add new stimuli sets that they wish to provide to the research community, we allow continuous content extension. Researchers who wish to add their new stimuli set to the searchable KAPODI database can fill out the corresponding set form available through https://linktr.ee/KAPODI_database. The submitted information will be uploaded to the database upon verification by one of the authors.


Declaration of Conflicting Interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iDs

Kathrin Diconne  <https://orcid.org/0000-0001-9568-2611>

Thomas J. Hostler  <https://orcid.org/0000-0002-4658-692X>

References

- Adolph, D., & Alpers, G. W. (2010). Valence and arousal: A comparison of two sets of emotional facial expressions. *American Journal of Psychology*, 123(2), 209–219. <https://doi.org/10.5406/amerjpsyc.123.2.0209>

- Barrington, B. L. (1963). A list of words descriptive of affective reactions. *Journal of Clinical Psychology*, 19(2), 259–262. [https://doi.org/10.1002/1097-4679\(196304\)19:2%3c259::AID-JCLP2270190238%3e3.0.CO;2-3](https://doi.org/10.1002/1097-4679(196304)19:2%3c259::AID-JCLP2270190238%3e3.0.CO;2-3)
- Bartlett, M. S., Hager, J. C., Ekman, P., & Sejnowski, T. J. (1999). Measuring facial expressions by computer image analysis. *Psychophysiology*, 36(2), 253–263. <https://doi.org/10.1017/S0048577299971664>
- Battocchi, A., Pianesi, F., & Goren-Bar, D. (2005). The properties of DaFEx, a database of kinetic facial expressions. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3784, 558–565. https://doi.org/10.1007/11573548_72
- Bertels, J., Deliens, G., Peigneux, P., & Destrebecqz, A. (2014). The Brussels Mood Inductive Audio Stories (MIAS) database. *Behavior Research Methods*, 46(4), 1098–1107. <https://doi.org/10.3758/s13428-014-0445-3>
- Boiger, M., Ceulemans, E., De Leersnyder, J., Uchida, Y., Norasakkunkit, V., & Mesquita, B. (2018). Beyond essentialism: Cultural differences in emotions revisited. *Emotion (Washington, D.C.)*, 18(8), 1142–1162. <https://doi.org/10.1037/emo0000390>
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49–59. [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*, Technical report C-1, The Center for Research in Psychophysiology, University of Florida, 30(1), 25–36.
- Calvo, M. G., Fernández-Martín, A., Recio, G., & Lundqvist, D. (2018). Human observers and automated assessment of dynamic emotional facial expressions: KDEF-dyn database validation. *Frontiers in Psychology*, 9, 1–12. <https://doi.org/10.3389/fpsyg.2018.02052>
- Cardos, R., Predatu, R., & David, O. (2017). Development and validation of a cartoon based set of children's facial emotion stimuli for the RETHink online therapeutic game. In *Cognitive Behavioral Coaching ICCBC 2016* (pp. 39–43). Medimond.
- Costantini, G., Iadarola, I., Paoloni, A., & Todisco, M. (2014). EMOVO corpus: An Italian emotional speech database. *Proceedings of the 9th International Conference on Language Resources and Evaluation*, 3501–3504.
- Crane, E., & Gross, M. (2007). Motion capture and emotion: Affect detection in whole body movement. In A. C. R. Paiva, R. Prada, & R. W. Picard (Eds.), *Affective computing and intelligent interaction. ACII 2007. Lecture notes in computer science* (Vol. 4738, pp. 95–101). Springer. https://doi.org/10.1007/978-3-540-74889-2_9
- Cullen, A., & Harte, N. (2018). A longitudinal database of Irish political speech with annotations of speaker ability. *Language Resources and Evaluation*, 52(2), 401–432. <https://doi.org/10.1007/s10579-017-9401-z>
- DeBusk, K. P. A., & Austin, E. J. (2011). Emotional intelligence and social perception. *Personality and Individual Differences*, 51(6), 764–768. <https://doi.org/10.1016/j.paid.2011.06.026>
- Delatorre, P., Salguero, A. G., León, C., & Tapscott, A. (2019). The impact of context on affective norms: A case of study with suspense. *Frontiers in Psychology*, 10, 1988. <https://doi.org/10.3389/fpsyg.2019.01988>
- Deng, Y., Yang, M., & Zhou, R. (2017). A new standardized emotional film database for Asian culture. *Frontiers in Psychology*, 8, 1941. <https://doi.org/10.3389/fpsyg.2017.01941>
- Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES-a database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, 42(1), 351–362. <https://doi.org/10.3758/BRM.42.1.351>
- Edelman, M., Bourdieu, P., Thompson, J. B., Raymond, G., & Adamson, M. (1992). Language and symbolic power. *Contemporary Sociology*, 21(5), 717. <https://doi.org/10.2307/2075589>
- Ekman, P. (1976). *Pictures of facial affect*. Consulting Psychologists Press.
- Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science (New York, N.Y.)*, 164(3875), 86–88. <https://doi.org/10.1126/science.164.3875.86>
- Ferré, P., Guasch, M., Moldovan, C., & Sánchez-Casas, R. (2012). Affective norms for 380 Spanish words belonging to three different semantic categories. *Behavior Research Methods*, 44(2), 395–403. <https://doi.org/10.3758/s13428-011-0165-x>
- Frijda, N. H. (2007). *The laws of emotion*. Lawrence Erlbaum.
- Grandy, T. H., Lindenberger, U., & Schmiedek, F. (2020). Vampires and nurses are rated differently by younger and older adults - Age-comparative norms of imageability and emotionality for about 2500 German nouns. *Behavior Research Methods*, 52(3), 980–989. <https://doi.org/10.3758/s13428-019-01294-2>
- Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, 26(1), 1–26. <https://doi.org/10.1080/1047840X.2014.940781>
- Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition and Emotion*, 9(1), 87–108. <https://doi.org/10.1080/02699939508408966>
- Grühn, D., & Sharifian, N. (2016). Lists of emotional stimuli. In *Emotion measurement* (pp. 145–164). Woodhead Publishing. <https://doi.org/10.1016/b978-0-08-100508-8.00007-2>
- Gutiérrez-Maldonado, J., Rus-Calafell, M., & González-Conde, J. (2014). Creation of a new set of dynamic virtual reality faces for the assessment and training of facial emotion recognition ability. *Virtual Reality*, 18(1), 61–71. <https://doi.org/10.1007/s10055-013-0236-7>
- Hall, J. A., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion (Washington, D.C.)*, 4(2), 201–206. <https://doi.org/10.1037/1528-3542.4.2.201>
- Haro, J., Ferré, P., Boada, R., & Demestre, J. (2017). Semantic ambiguity norms for 530 Spanish words. *Applied Psycholinguistics*, 38(2), 457–475. <https://doi.org/10.1017/S0142716416000266>
- Hayes, M. H., & Patterson, D. G. (1921). Experimental development of the graphic rating method. *Psychological Bulletin*, 18(1), 98–99.
- Hewig, J., Hagemann, D., Seifert, J., Gollwitzer, M., Naumann, E., & Bartussek, D. (2005). A revised film set for the induction of basic emotions. *Cognition and Emotion*, 19(7), 1095–1109. <https://doi.org/10.1080/02699930541000084>
- Holland, C. A. C., Ebner, N. C., Lin, T., & Samanez-Larkin, G. R. (2019). Emotion identification across adulthood using the dynamic FACES database of emotional expressions in younger, middle aged, and older adults. *Cognition and Emotion*, 33(2), 245–257. <https://doi.org/10.1080/02699931.2018.1445981>
- Imbir, K. K. (2015). Affective norms for 1,586 polish words (ANPW): Duality-of-mind approach. *Behavior Research Methods*, 47(3), 860–870. <https://doi.org/10.3758/s13428-014-0509-4>
- Imbir, K. K. (2016). Affective norms for 718 polish short texts (ANPST): Dataset with affective ratings for valence, arousal, dominance, origin, subjective significance and source dimensions. *Frontiers in Psychology*, 7, 1030. <https://doi.org/10.3389/fpsyg.2016.01030>
- Isaacowitz, D. M., Löckenhoff, C. E., Lane, R. D., Wright, R., Sechrest, L., Riedel, R., & Costa, P. T. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging*, 22(1), 147–159. <https://doi.org/10.1037/0882-7974.22.1.147>
- Izard, C. E. (2007). Basic emotions, natural kinds. *Emotion Schemas, and a New Paradigm. Perspectives on Psychological Science*, 2(3), 260–280. <https://doi.org/10.1111/j.1745-6916.2007.00044.x>
- Kanske, P., & Kotz, S. A. (2011). Cross-modal validation of the Leipzig Affective Norms for German (LANG). *Behavior Research Methods*, 43(2), 409–413. <https://doi.org/10.3758/s13428-010-0048-6>
- Kanske, P., & Kotz, S. A. (2012). Auditory affective norms for German: Testing the influence of depression and anxiety on valence and arousal ratings. *PLoS ONE*, 7(1), 1–6. <https://doi.org/10.1371/journal.pone.0030086>
- Katsimerou, C., Albeda, J., Hultgren, A., Heynderickx, I., & Redi, J. A. (2016). Crowdsourcing empathetic intelligence: The case of the annotation of EMMA database for emotion and mood recognition. *ACM Transactions on Intelligent Systems and Technology*, 7(4), 1–27. <https://doi.org/10.1145/2897369>

- Kleinsmith, A., Bianchi-Berthouze, N., & Steed, A. (2011). Automatic recognition of non-acted affective postures. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 41(4), 1027–1038. <https://doi.org/10.1109/tsmcb.2010.2103557>
- Kring, A. M., & Gordon, A. H. (1998). Sex differences in emotion: Expression, experience, and physiology. *Journal of Personality and Social Psychology*, 74(3), 686–703. <https://doi.org/10.1037/0022-3514.74.3.686>
- Krumhuber, E. G., Skora, L., Küster, D., & Fou, L. (2017). A review of dynamic datasets for facial expression research. *Emotion Review*, 9(3), 280–292. <https://doi.org/10.1177/1754073916670022>
- Lang, P. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 119–137). Ablex. <https://doi.org/10.1111/j.1469-8986.1993.tb03352.x>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, 1(39–58), 3.
- Lassalle, A., Pigat, D., O'Reilly, H., Berggen, S., Fridenson-Hayo, S., Tal, S., Elfström, S., Råde, A., Golan, O., Bölte, S., Baron-Cohen, S., & Lundqvist, D. (2019). The EU-emotion voice database. *Behavior Research Methods*, 51(2), 493–506. <https://doi.org/10.3758/s13428-018-1048-1>
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140), 5–55.
- Lithari, C., Frantzidis, C. A., Papadelis, C., Vivas, A. B., Klados, M. A., Kourtidou-Papadeli, C., Pappas, C., Ioannides, A. A., & Bamidis, P. D. (2010). Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions. *Brain Topography*, 23(1), 27–40. <https://doi.org/10.1007/s10548-009-0130-5>
- Little, A. C. (2013). The influence of steroid sex hormones on the cognitive and emotional processing of visual stimuli in humans. *Frontiers in Neuroendocrinology*, 34(4), 315–328. <https://doi.org/10.1016/j.yfine.2013.07.009>
- LoBue, V., & Thrasher, C. (2015). The Child Affective Facial Expression (CAFE) set: Validity and reliability from untrained adults. *Frontiers in Psychology*, 5, 1532. <https://doi.org/10.3389/fpsyg.2014.01532>
- Marcell, M. M., Borella, D., Greene, M., Kerr, E., & Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of Clinical and Experimental Neuropsychology*, 22(6), 830–864. <https://doi.org/10.1076/jcen.22.6.830.949>
- Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology*, 1(1), 167–195. <https://doi.org/10.1146/annurev.clinpsy.1.102803.143916>
- Matsumoto, D., Yoo, S. H., Nakagawa, S., Alexandre, J., Altarriba, J., Anguas-Wong, A. M., Arriola, M., Bauer, L. M., Bond, M. H., Cabecinhas, R., Chae, J., Comunian, A. L., DeGere, D. N., de Melo Garcia Bley, L., Fok, H. K., Friedlmeier, W., Garcia, F. M., Ghosh, A., & Granskaya, J. V., ... Yoo, S. H. (2008). Culture, emotion regulation, and adjustment. *Journal of Personality and Social Psychology*, 94(6), 925–937. <https://doi.org/10.1037/0022-3514.94.6.925>
- McDuff, D., Amr, M., & Kaliouby, R. E. (2019). AM-FED +: An extended dataset of naturalistic facial expressions collected in everyday settings. *IEEE Transactions on Affective Computing*, 10(1), 7–17. <https://doi.org/10.1109/TAFFC.2018.2801311>
- Miccoli, L., Delgado, R., Guerra, P., Versace, F., Rodríguez-Ruiz, S., & Fernández-Santaella, M. C. (2016). Affective pictures and the open library of affective foods (OLAF): Tools to investigate emotions toward food in adults. *PLoS ONE*, 11(8), 1–13. <https://doi.org/10.1371/journal.pone.0158991>
- Miccoli, L., Delgado, R., Rodríguez-Ruiz, S., Guerra, P., García-Mármol, E., & Fernández-Santaella, F. (2014). Meet OLAF, a good friend of the IAPS! the open library of affective foods: A tool to investigate the emotional impact of food in adolescents. *PLoS ONE*, 9(12), 1–22. <https://doi.org/10.1371/journal.pone.0114515>
- Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, 36(4), 630–633. <https://doi.org/10.3758/bf03206543>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. <https://doi.org/10.1136/bmj.b2535>
- Montefinese, M., Ambrosini, E., Fairfield, B., & Mammarella, N. (2014). The adaptation of the Affective Norms for English Words (ANEW) for Italian. *Behavior Research Methods*, 46(3), 887–903. <https://doi.org/10.3758/s13428-013-0405-3>
- Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Du Sert, N. P., Simonsohn, U., Wagenmakers, E. J., Ware, J. J., & Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1), 1–9. <https://doi.org/10.3758/s13428-013-0405-3>
- Nater, U. M., Abbruzzese, E., Krebs, M., & Ehrlert, U. (2006). Sex differences in emotional and psychophysiological responses to musical stimuli. *International Journal of Psychophysiology*, 62(2), 300–308. <https://doi.org/10.1016/j.ijpsycho.2006.05.011>
- Neiberg, D., & Elenius, K. (2008). Automatic recognition of anger in spontaneous speech. *Proceedings of the Annual Conference of the International Speech Communication Association, Interspeech*, 2755–2758.
- Nieuwenhuis-Mark, R. E., Schalk, K., & De Graaf, N. (2009). Free recall and learning of emotional word lists in very elderly people with and without dementia. *American Journal of Alzheimer's Disease and Other Dementias*, 24(2), 155–162. <https://doi.org/10.1177/1533317508330561>
- Okon-Singer, H., Kofman, O., Tzelgov, J., & Henik, A. (2011). Using international emotional picture sets in countries suffering from violence. *Journal of Traumatic Stress*, 24(2), 239–242. <https://doi.org/10.1002/jts.20600>
- Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, 49(3), 197. <https://doi.org/10.1037/h0055737>
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning* (No. 47). University of Illinois press. <https://doi.org/10.1017/s0008413100018740>
- O'Toole, A. J., Harms, J., Snow, S. L., Hurst, D. R., Pappas, M. R., Ayyad, J. H., & Abdi, H. (2005). A video database of moving faces and people. *IEEE Transactions on pattern analysis and machine intelligence*, 27(5), 812–816. <https://doi.org/10.1109/TPAMI.2005.90>
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), 210. <https://doi.org/10.1186/s13643-016-0384-4>
- Parsons, C. E., Young, K. S., Craske, M. G., Stein, A. L., & Kringelbach, M. L. (2014). Introducing the Oxford Vocal (OxVoc) sounds database: A validated set of non-acted affective sounds from human infants, adults, and domestic animals. *Frontiers in Psychology*, 5, 562. <https://doi.org/10.3389/fpsyg.2014.00562>
- Pollak, S. D., Camras, L. A., & Cole, P. M. (2019). Progress in understanding the emergence of human emotion. *Developmental Psychology*, 55(9), 1801–1811. <https://doi.org/10.1037/dev0000789>
- Prada, M., Garrido, M. V., Camilo, C., & Rodrigues, D. L. (2018). Subjective ratings and emotional recognition of children's facial expressions from the CAFE set. *PLoS ONE*, 13(12), e0209644. <https://doi.org/10.1371/journal.pone.0209644>
- Pronk, T., van Deursen, D. S., Beraha, E. M., Larsen, H., & Wiers, R. W. (2015). Validation of the Amsterdam beverage picture set: A controlled picture set for cognitive bias measurement and modification paradigms. *Alcoholism: Clinical and Experimental Research*, 39(10), 2047–2055. <https://doi.org/10.1111/acer.12853>
- Provost, E. M., Shangguan, Y., & Busso, C. (2015). UMEME: University of Michigan emotional mcgurr effect data set. *IEEE Transactions on Affective Computing*, 6(4), 395–409. <https://doi.org/10.1109/TAFFC.2015.2407898>
- Ramalingam, V. V., Pandian, A., Jaiswal, A., & Bhatia, N. (2018). Emotion detection from text. *Journal of Physics: Conference Series*, 1000(1), 012027. <https://doi.org/10.1088/1742-6596/1000/1/012027>

- Ribeiro, R. L., Pompéia, S., & Bueno, O. F. A. (2005). Comparison of Brazilian and American norms for the International Affective Picture System (IAPS). *Brazilian Journal of Psychiatry*, 27(3), 208–215. <https://doi.org/10.1590/S1516-44462005000300009>
- Riegel, M., Wierzbna, M., Wypych, M., Żurawski, Ł., Jednoróg, K., Grabowska, A., & Marchewka, A. (2015). Nencki affective word list (NAWL): The cultural adaptation of the Berlin Affective Word List–Reloaded (BAWL-R) for Polish. *Behavior Research Methods*, 47(4), 1222–1236. <https://doi.org/10.3758/s13428-014-0552-1>
- Sacco, A. M., de Paula Couto, M. C. P., & Koller, S. H. (2016). Construction and validation of the White, Pardo, and Black children picture set (BIC-multicolor). *Psychology and Neuroscience*, 9(1), 68–78. <https://doi.org/10.1037/pne0000040>
- Samaria, F. S., & Harter, A. C. (1994). Parameterisation of a stochastic model for human face identification. *IEEE Workshop on Applications of Computer Vision - Proceedings*, 138–142. <https://doi.org/10.1109/acv.1994.341300>
- Schmidtke, D. S., Schröder, T., Jacobs, A. M., & Conrad, M. (2014). ANGST: Affective norms for German sentiment terms, derived from the affective norms for English words. *Behavior Research Methods*, 46(4), 1108–1118. <https://doi.org/10.3758/s13428-013-0426-y>
- Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, 106(2), 975–983. <https://doi.org/10.1016/j.cognition.2007.03.005>
- Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2019). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*, 51(3), 1258–1270. <https://doi.org/10.3758/s13428-018-1099-3>
- Shankland, R., Favre, P., Corubolo, D., Méary, D., Flaudias, V., & Mermillod, M. (2019). Food-Cal: Development of a controlled database of high and low calorie food matched with non-food pictures. *Eating and Weight Disorders*, 24(6), 1041–1050. <https://doi.org/10.1007/s40519-019-00687-8>
- Sloan, D. M., Sege, C. T., McSweeney, L. B., Suvak, M. K., Shea, M. T., & Litz, B. T. (2010). Development of a borderline personality disorder-relevant picture stimulus set. *Journal of Personality Disorders*, 24(5), 664–675. <https://doi.org/10.1521/pedi.2010.24.5.664>
- Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods*, 44(1), 256–269. <https://doi.org/10.3758/s13428-011-0131-7>
- Soleymani, M., Lichtenauer, J., Pun, T., & Pantic, M. (2012). A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing*, 3(1), 42–55. <https://doi.org/10.1109/T-AFFC.2011.25>
- Sutton, T. M., & Altarriba, J. (2016). Color associations to emotion and emotion-laden words: A collection of norms for stimulus construction and selection. *Behavior Research Methods*, 48(2), 686–728. <https://doi.org/10.3758/s13428-015-0598-8>
- Szwoch, M. (2014). On Facial Expressions and Emotions RGB-D Database. In: Kozielski S., Mrozek D., Kasprowski P., Małysiak-Mrozek B., Kostrzewa D. (eds) *Beyond Databases, Architectures, and Structures. BDAS 2014. Communications in Computer and Information Science*, vol 424. Springer, Cham. https://doi.org/10.1007/978-3-319-06932-6_37
- Szymanska, M., Comte, A., Tio, G., Vidal, C., Monnin, J., Smith, C. C., Nezelof, S., & Vulliez-Coady, L. (2019). The Besançon affective picture set-adult (BAPS-Adult): Development and validation. *Psychiatry Research*, 271(2019), 31–38. <https://doi.org/10.1016/j.psychres.2018.11.005>
- Szymanska, M., Monnin, J., Noiret, N., Tio, G., Galdon, L., Laurent, E., Nezelof, S., & Vulliez-Coady, L. (2015). The besançon affective picture set-adolescents (the BAPS-Ado): Development and validation. *Psychiatry Research*, 228(3), 576–584. <https://doi.org/10.1016/j.psychres.2015.04.055>
- Thayer, S. (1980a). The effect of expression sequence and expressor identity on judgments of the intensity of facial expression. *Journal of Nonverbal Behavior*, 5(2), 71–79. <https://doi.org/10.1007/BF00986510>
- Thayer, S. (1980b). The effect of facial expression sequence upon judgments of emotion. *The Journal of Social Psychology*, 111(2), 305–306.
- Torkamani-Azar, M., Kanik, S. D., Vardan, A. T., Aydin, C., & Cetin, M. (2019). Emotionality of Turkish language and primary adaptation of affective English norms for Turkish. *Current Psychology*, 38(2), 273–294. <https://doi.org/10.1007/s12144-018-0119-x>
- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, 168(3), 242–249. <https://doi.org/10.1016/j.psychres.2008.05.006>
- Tu, Y. Z., Lin, D. W., Suzuki, A., & Goh, J. O. S. (2018). East Asian young and older adult perceptions of emotional faces from an age-and sex-fair east Asian facial expression database. *Frontiers in Psychology*, 9, 23. <https://doi.org/10.3389/fpsyg.2018.02358>
- Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, 26(4), 557–580. [https://doi.org/10.1002/\(sici\)1099-0992\(199607\)26:4<557::aid-ejsp769>3.0.co;2-4](https://doi.org/10.1002/(sici)1099-0992(199607)26:4<557::aid-ejsp769>3.0.co;2-4)
- Wingenbach, T. S., Ashwin, C., & Brosnan, M. (2016). 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*, 11(1), e0147112. <https://doi.org/10.1371/journal.pone.0147112>
- Yan, W. J., Wu, Q., Liu, Y. J., Wang, S. J., & Fu, X. (2013). CASME database: A dataset of spontaneous micro-expressions collected from neutralized faces. *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, FG 2013*. <https://doi.org/10.1109/FG.2013.6553799>
- Yingliang, M. A., Paterson, H. M., & Pollick, F. E. (2006). A motion capture library for the study of identity, gender, and emotion perception from biological motion. *Behavior Research Methods*, 38(1), 134–141. <https://doi.org/10.3758/bf03192758>
- Zafeiriou, S., Papaioannou, A., Kotsia, I., Nicolaou, M., & Zhao, G. (2016). Facial affect “In-the-Wild”: A survey and a new database. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 1487–1498. <https://doi.org/10.1109/CVPRW.2016.186>
- Zander, M. F. (2006). Musical influences in advertising: How music modifies first impressions of product endorsers and brands. *Psychology of Music*, 34(4), 465–480. <https://doi.org/10.1177/0305735606067158>