Tutorial on Designing, Implementing, and Analyzing a Degraded Image Paradigm: A Facial Expression-Decoding Task Example

Liana S. E. Hone, Ph.D.[a,b,1] (orcid.org/0000-0002-6777-978)

John E. Scofield, M.S.[a] (orcid.org/0000-0001-6345-1181)

Bruce D. Bartholow, Ph.D.[a] (orcid.org/0000-0002-9234-6417)

David C. Geary, Ph.D.[a,c] (orcid.org/0000-0003-3029-6343)

[a]Department of Psychological Sciences, University of Missouri

[b]Clinical and Research Institute on Addictions, University at Buffalo

[c]Interdisciplinary Neuroscience Program, University of Missouri

Author Note

This work was supported by grants F32AA025830 (PI: Hone) and R01AA020970 (PI: Bartholow) from the National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health.

[1]Corresponding author: 1021 Main St., Buffalo, NY 14203; Phone: 716-887-2242;

E-mail: lhone@buffalo.edu.

Abstract

Although various versions of degraded image tasks exist, the complete process of creating these tasks can be effortful, time-consuming, and costly. The aim of the present tutorial is to provide a step-by-step, user-friendly, economical guide to creating a degraded image task using freely available resources, and provide resources to create multiple experiments adaptable to address many research questions. We demonstrated the utility of this method by designing, implementing, and analyzing a facial expression-decoding task, and confirmed the validity of the task by replicating often-found sex differences in face processing. Using this tutorial will enable new investigators with limited prior experience and available funding to quickly and easily design experiments using degraded image tasks, thereby reducing the delay between idea conceptualization and data acquisition.

Keywords: tutorial, degraded image, facial expression-decoding, GIMP, PsychoPy, E-Prime, R

Introduction

Social acuity, or the ability to decode nonverbal behavior, including facial expressions, is a critical component of human communication and thus has been a topic of continued interest among scientists for more than a century (Darwin, 1872; Ghiselin, Ekman, & Gruber, 1974). The ability to identify facial expressions is thought to be an adaptive human universal (Brown, 1991; Ekman, 1994; Ekman & Friesen, 1971; Izard, 1994; c.f. Russell, 1994). Indeed, neuroimaging studies suggest there are distinct face emotion processing areas (Matsuda et al., 2013), with lesions in these areas causing deficits in recognizing emotions signaled by facial expressions (Adolphs, Baron-Cohen, & Tranel, 2002; Adolphs, Tranel, Damasio, & Damasio, 1994; Eslinger & Damasio, 1985; Stone, Baron-Cohen, Calder, Keane, & Young, 2003; Wager, Phan, Liberzon, & Taylor, 2003); see Rossion, (2014) for a general review of understanding face processing.

Given the key role of communication, cooperation, and sociality in our species (Trivers, 1971), it makes sense that the ability to identify facial expressions has numerous implications for health and well-being (Carton, Kessler, & Pape, 1999). Social acuity impairments are related to, for example, autism (Harms, Martin, & Wallace, 2010), aging (Ruffman, Henry, Livingstone, & Phillips, 2008), borderline personality disorder (Domes, Schulze, & Herpertz, 2009), anorexia (Oldershaw et al., 2011), and schizophrenia (Edwards, Jackson, & Pattison, 2002). Thus, studies of facial expression-decoding are relevant to neuropsychology, clinical psychology, developmental psychology, experimental psychology, cognitive science, behavioral neuroscience, and social psychology.

Numerous experimental paradigms are available to measure social acuity, each with their strengths and weaknesses (Elfenbein, Marsh, & Ambady, 2002; Wilhelm, Hildebrandt, Manske, Schacht, & Sommer, 2014). Humans' ability to decode facial expressions has been assessed in hundreds of studies using these various paradigms (Hess, Blairy, & Kleck, 1997; Tottenham et al., 2009). Some simple paradigms merely require participants to view a collection of photographs of faces depicting static emotions, and to select from a list of emotion words the one that best describes the face in the photograph (e.g., Emotion Recognition Experiment; Izard, 1971). There are several variations on this type of task, including the Facially Expressed Emotion Labeling (FEEL) task (Kessler, Bayerl, Deighton, & Traue, 2002) and other simple tasks (Pollatos, Herbert, Schandry, & Gramann, 2008) requiring participants to categorize emotions depicted in images from freely available face databases (Lundqvist, Flykt, & Öhman, 1998).

More complicated tasks have been developed as well, including those that are composed of morphed faces. These tasks typically involve the display of neutral faces (0% emotion), followed by faces depicting increasing levels of emotion intensity, and then finally display the full intensity of the emotion (100% emotion). Some versions of morphed-faces tasks involve the display of emotions on a continuum, for example, morphing from happy to sad, rather than morphing from neutral to intense emotion. Wilhelm and colleagues (2014) provide a useful discussion of available tasks for measuring emotion perception and recognition from faces, including their shortcomings (e.g., psychometrics and task applicability), as well as a 16-task test battery for measuring facial expression-decoding.

Depending on a user's goal, facial expression-decoding paradigms could benefit from the use of degraded images (Dolan et al., 1997; Sadr & Sinha, 2004). Degraded image tasks are exceedingly useful in psychology broadly, not just in the study of social acuity. For example, Snodgrass, Smith, Feenan, and Corwin (1987) developed a computerized degraded image task based on the Gollin Picture Test (Gollin, 1960) to examine perceptual learning. Macrae, Stangor, and Milne (1994) used degraded images of stereotype-relevant words to show that priming with a stereotype enhances word recognition. Similarly, Fazio, Williams, and Powell (2000) used the extent to which a category name facilitated recognition of a degraded item as a measure of associative strength in memory. James, Goodale, Menon, Humphrey, and Gati (2000) created a task in which pixels of noise were gradually removed to reveal an image in order to slow down the process of recognition during their study of brain activation. Additionally, Eberhardt, Goff, Purdie, and Davies (2004) used degraded images of guns to examine the extent to which activating the Black racial category lowered the perceptual threshold for recognizing guns.

With regard to the use of degraded images in the study of face perception, Grady and coleagues (1996) used a face matching task with static degraded images to study age-related changes in cerebral blood flow during visuoperceptual processing. More recently, Sadr and Sinha (2004) created a task in which objects, including faces, evolve from and then dissolve back into randomness, and discuss how to preserve important low-level properties of an image using the Random Image Structure Evolution (RISE) technique. Additionally, Rossion and Caharel (2011) used degraded images of neutral faces and cars to study how quickly stimuli are categorized as faces and Rossion, Hanseeuw, and Dricot (2012) used these same degraded images to study face perception areas in the brain. Royer, Blais, Gosselin, Duncan, and Fiset (2015) used degraded images of neutral faces to show that stimuli involving the whole face are not required in face processing—some individuals can identify faces with degraded visual information. Finally, Wentura and Rohr (2018) used priming methods to study whether relevance, arousal, or specificity are important in differentiation of static degraded images of emotional faces.

Though many studies have used degraded face images in face perception and neuroimaging studies, to our knowledge no studies have yet used degraded images of faces with pixelated noise that slowly come into focus in the context of facial expression-decoding. Available paradigms used to study decoding of facial expressions evince several strengths, including construct, predictive, and face validity. However, a more flexible paradigm that includes the use of degraded images could address several limitations of existing methods. First, many existing tasks are limited by the use of a single face database or a database that does not meet the needs of the experiment at hand. A researcher may want to vary the sex, race, or age of the faces used in expression-decoding tasks, vary whether the mouth is open or closed, or vary the intensity of the expression displayed (Tottenham et al., 2009). Furthermore, new face databases with more detail are now available, including the Multi-Racial Mega-Resolution database comprising 74 extremely high resolution images of European, African, and East Asian faces (Strohminger et al., 2015). A flexible task would allow for ease of incorporating novel, diverse, naturalistic, high-resolution, and controlled face images from multiple datasets in future studies. A flexible task would also allow for the use of face-like images, distorted face images, or images of faces that have been combined or edited to serve the needs of the experiment at hand.

A second limitation of traditional social acuity tasks that use morphed faces is that they may seem unnatural (Schultz & Pilz, 2009). This is because the process of creating videos from multiple frames of static pictures, showing a progression of a neutral face to a face depicting the full intensity of an emotion, can cause peculiar blending artifacts to appear. A degraded image paradigm avoids this issue by allowing participants to see faces displaying various emotion intensities come slowly into focus. Finally and most importantly, most existing facial expression-decoding tasks are costly to access, limiting their availability to many researchers.

Given these limitations, it can be beneficial for researchers to design their own tasks. Yet, doing so comes with the costs of implementing and analyzing data from new paradigms.

Therefore, the purpose of this tutorial is to demonstrate the creation of stimuli, along with detailed information on how to implement and analyze experiments using degraded image tasks following the principles outlined in previous facial expression-decoding paradigms. Consequently, our tutorial provides methodological advances over prior research. This tool has been successfully implemented to create a facial expression-decoding task with face stimuli matching the demographic of a sample of undergraduates at a large Midwestern University (i.e., predominately White). We use this as an example degraded image task below and provide access to this task on the Open Science Framework (OSF). All steps use freely available resources.

Description, Technical Specifications, and Availability

This tutorial provides and describes how to use freely available image databases, software, and statistical packages to degrade images, present stimuli, and extract data in new experiments, using Windows 7 or later. The images used in this degraded image paradigm were faces selected from the MMI Facial Expression Database (Pantic, Valstar, Rademaker, & Maat, 2005; Valstar & Pantic, 2010), the Karolinska Directed Emotional Faces (KDEF) database (Lundqvist et al., 1998), the Radboud Faces Database (RaFD; Langner et al., 2010), and the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015; see Table 1). Ma and colleagues (2015) provide an overview of face databases as well as the strengths and weaknesses of each. The MMI Facial Expression Database (https://mmifacedb.eu/) comprises more than 2900 high-resolution still images of 75 individuals. The database is freely available and any reports on research that use the MMI Facial Expression Database are required to acknowledge the MMI Facial Expression Database in the following way: “Portions of the research in this paper uses the MMI Facial Expression Database collected by Valstar and Pantic.” Researchers should send a copy of any reports to the authors of the MMI Facial Expression Database (Pantic et al., 2005).

The Karolinska Directed Emotional Faces (KDEF; http://kdef.se/index.html) comprises 4900 pictures of 70 individuals. The database is freely available and any reports on research that use the KDEF database should cite Lundqvist and colleagues (1998). The Radboud Faces Database (RaFD; http://www.socsci.ru.nl/rafd) comprises pictures of 67 individuals displaying 8 emotional expressions. The database is freely available and any reports on research that use the RaFD should cite Langner and colleagues (2010). The Chicago Face Database comprises pictures of 158 Black and White males and females between the ages of 17-65. The database is freely available for non-commercial use and should acknowledge Ma and colleagues (2015).

Images of faces were degraded by adding pixelated noise and exported as stimulus movie files using GNU Image Manipulation (GIMP) version 2.10.8 (https://www.gimp.org), a freely available, cross-platform image editor (The GIMP Team, 2019a). The version of GIMP used in conjunction with this tutorial is available on OSF. We provide an interactive video tutorial on how to degrade any image by adding pixelated noise and how to create stimulus movie files that slowly remove pixelated noise from the degraded images. This component is especially useful for researchers who would like to expand the task to include additional faces varied on sex, age, race, and expression, or adapt the program to display any number of other degraded objects or words.

The stimulus movie files can be presented using PsychoPy3 version 3.1.5 (https://www.psychopy.org), a freely available, open-source software tool for controlling stimulus presentation and recording responses (Peirce et al., 2019), or E-Prime 3 (https://pstnet.com/products/e-prime), which is not freely available but is widely used in psychological laboratories (Psychology Software Tools, 2016). E-Prime recommends that researchers use a millisecond-accurate stimulus control and response input system, such as

Chronos (Psychology Software Tools, 2015). Using Chronos (or a similar device) ensures that stimulus and response timing is accurate to within 1-2 milliseconds (Psychology Software Tools, 2015). In the Example described below, we used E-Prime. We provide access to already programmed experimental PsychoPy and E-Prime files that present stimulus movie files and record behavioral responses.

Data were analyze using R version 3.5.1 (https://www.r-project.org), a freely available, cross-platform software environment for statistical computing and graphics. We provide instructions for using R to extract data from both PsychoPy and E-Prime, and provide to R code for loading of data into a format useable for statistical analyses. See Figure 1 for an overview of the tutorial and Table 2 for a list of all software used in the tutorial.

Designing New Experiments

Degrade Images by Adding Pixelated Noise

We used the principles outlined by (James and colleagues (2000) to create a task in which pixels of noise were gradually removed to reveal an image. We created 6000mpb videos with scrambled pixels that slowly come into focus. To create these videos, we used a predefined function in GIMP. We opened images saved from face databases in GIMP (The GIMP Team, 2019b). We selected “filters,” then “noise,” then “spread.” We increased the amount of spread in horizontal and vertical boxes to match the largest dimension of the image (e.g., if the image was 438 x 435, we increased spread amount to 438). The user can find the dimensions of an image by right clicking on the image, selecting “properties,” and then “details.” After the filter was applied to the image, we saved the image as a new file and closed the image with the spread filter applied. We then opened the original image again. We selected “layers” in the top left tool bar, and then clicked the “duplicate layer” button (Shift+Crtl+D) 77 times.

At the bottom of the tool bar, we then selected the original image and changed the "opacity" to zero. We then selected the original copy layer in the bottom of the layer toolbar and changed the opacity to zero. We increased each image above the original image by an opacity of 1.3 until we reached an opacity of 100 (i.e., Copy = opacity 1.3, Copy #1 = opacity 2.6, Copy #2 = 3.9, Copy #3 = opacity 5.2, and so on). We then selected "file," and "open as layers," and selected the image with the "spread" filter previously saved. We duplicated the "spread" image layer (Shift+Crtl+D) and then moved it below the original image layer at the bottom of the toolbar. We repeated this step until there was a spread image layer below each copy layer of the original image.

We then selected the copy of the original image layer with the opacity of 100. Then we selected "layer" and then "merge down." You will know you did this correctly if the original copy image and the spread image underneath it is labeled, "*Spread.xcf Copy #…*" We repeated this step until all of the original image layers were merged down into the spread image layers. When you scroll through the tool bar, the top images should be clear, the middle images should be progressively blurrier, and the bottom images should be blurry. This approach introduces noise as a spatial filter across an image. The rate at which noise is removed from the pixelated image is constant across time. A video depicting this process is available on OSF.

Create Stimulus Movie Files

To create the movie files for each stimulus, we then selected "export as." Under "name," we changed *.png* to *.gif* in the file name box, and then selected "export." In the resulting box with "gif options," we selected the box titled "as animation." We changed the delay between the 78 frames to 600 milliseconds and selected "export." This resulted in stimulus movie files that were approximately 45 seconds long. This can easily be adjusted to accommodate speeded decision paradigms with a range of times. A video depicting this process is available on OSF. Depending on the user's goals, we suggest exporting stimulus movie files as *.avi*, *.mov*, or *.mp4* files instead of *.gif* files as *.gif* files may have temporal lags. If *.gif* files are used, actual presentation time of the stimulus movie files should be recorded and compared to the expected 45 seconds.

Use PsychoPy or E-Prime to Present Stimuli for Data Collection

Once stimulus movie files are created, they can be loaded into programs for viewing and data collection. The PsychoPy and E-Prime programs are commonly used in experimental psychology for stimulus presentation and response acquisition. An in-depth tutorial on how to use these programs is beyond this tutorial, but 36 stimulus movie files created using the instructions above, 34 of which are used in the "Example" section below, are freely available on OSF and are loaded into experimental files and ready to be used in both PsychoPy and E-Prime. If researchers wish to create their own stimuli using the guide discussed above, stimulus movie files can easily be replaced in the experimental files. Instruction documents for editing the "Example" task can be found alongside each of the PsychoPy and E-prime folders on OSF. Peirce and colleagues (2019) provide an overview on PsychoPy, and Richard and Charbonneau (2009) provide an overview on E-Prime.

The paradigm requires the participant to view a specified number of stimulus movie files (e.g., 34 stimulus movie files in the "Example") that present degraded, pixelated images on a computer monitor connected to a response input system (e.g., a keyboard). In each trial, a single image, randomly selected from the set of images, slowly comes into focus over the course of a specified duration of time (e.g., 45 seconds). Participants are required to halt the trial by pressing a button (e.g., the space bar) as soon as they recognize the image. A mask (e.g., *mask.jpg*, see OSF) is presented for a specified amount of time (e.g., 500 *ms*). This backward masking is commonly used to interfere with ongoing processing of stimuli (Del Zotto & Pegna, 2015; Koster, Verschuere, Burssens, Custers, & Crombez, 2007). A screen with response options is automatically displayed after the mask (this feature can be edited to suit the needs of the user). The trial ends when the participant selects a response option (e.g., by pressing the numeric key that corresponded to the number next to the chosen response). The next trial begins immediately after a response is recorded. The duration of intertrial interval can be varied according to the needs of a particular experiment. Primary dependent variables in the task include accuracy and reaction time.

Extract, Convert, and Aggregate Data Using R

The commands to read-in the data saved from PsychoPy and E-prime initially are different. PsychoPy saves the data in Excel format (*.csv*), whereas E-prime saves the data in a log file format (*.txt*). As PsychoPy automatically saves data in *.csv* format, datasets can be loaded using base R functions. You will simply set your working directory, and read in the dataset. With the preloaded “Example” experiment, the first three rows in the dataset correspond to instruction screens, and can be discarded upon loading the data into R. Code to perform this operation will look like this:

```
setwd("C:/put/directory/here/psychopy_experiment/data")
#load data file
data <- read.csv('psychopy_data.csv')
#first three rows are related to instructions and can be deleted
data <- data[-c(1:3), ]
```

If you are using E-prime, you will first need to install the R package *‘rprime’* (Mahr, 2015). Once installed, you can then load the package, set your working directory to the folder containing the data, and then load the data into R. Code to perform this operation will look like this:

```
install.packages('rprime')
library(rprime)
#set your working directory
setwd("C:/put/directory/here/eprime_files/")
#load data
data <- to_data_frame(FrameList(read_eprime("eprime_data.txt", remove_clock = T)))
```

Both E-prime and PsychoPy record participants' response times on each trial in milliseconds. Psychophy automatically scores response accuracy on each trial. However, when using E-prime, response accuracy will need to be manually calculated. Code for comparing the response on each trial with the correct response and saving binary accuracy to a new column in the same dataframe will look like this:

```
data$new_col <- ifelse(is.na(data$response), 0,
                    ifelse(data$correct == data$response, 1, 0))
```

A more concrete example of how to load, format, and analyze data using this task, especially in the case of loading and merging data from multiple participants, is available on OSF.

Example

Introduction

Sex differences in various social-cognitive competencies are well-documented and include an advantage of girls and women over boys and men in language fluency, decoding nonverbal cues (e.g., body posture), and sensitivity to subtle change in facial expression, among other advantages (Buck, Savin, Miller, & Caul, 1972; J. A. Hall & Matsumoto, 2004; McClure, 2000; Rosenthal, Archer, Hall, DiMatteo, & Rogers, 1979; van Beek & Dubas, 2008). Sex differences in these types of tasks range from modest (*d* ~ 0.25) to large (*d* > 1.0) depending on task difficulty (Rosenthal et al., 1979). There are different explanations for the source of these sex differences (Geary, Winegard, & Winegard, 2014; Kret & De Gelder, 2012), but these are not critical here. The critical finding is that women's advantage for the speed and accuracy of detecting subtle facial expressions ranges from *d* = 0.29 to 0.94 (J. K. Hall, Hutton, & Morgan, 2010; Hoffmann, Kessler, Eppel, Rukavina, & Traue, 2010; Sasson et al., 2010). The direction and magnitude of these sex differences provide a benchmark for evaluating the external validity of the image degradation procedure. External validity, the generalizability of results from one sample or context to another, can be achieved with systematic replication of studies—for example, studies of sex differences in facial-expression decoding—across samples and contexts using different methods to establish generalizable results (McDermott, 2011).

Method

Participants. We recruited 228 women (Age: *M* = 18.79, *SD* = 0.97) and 192 men (Age: *M* = 19.33, *SD* = 1.52) enrolled in introductory Psychology courses at a large Midwestern University to participate in a laboratory study of cognitive competencies, including our new facial expression-decoding task. The study was approved by an Institutional Review Board and all participants provided informed consent.

Materials. The images we selected to degrade were 34 faces (17 male; 17 female) each displaying one of seven emotions (happy, angry, surprise, fear, sad, disgust, neutral). We selected images so that our task displayed each of the seven emotions at least four times across the 34 trials, with at least two male faces and at least two female faces displaying each emotion (see Table 3). The faces were selected from freely available face image databases described above. We degraded all images by adding pixelated noise and created movie stimulus files using the instructions above. The degraded images that made up the stimulus movie files comprised 78 frames that range from completely obscured to completely clear with regard to pixelated noise. The delay between the 78 frames was 600 milliseconds which yielded stimulus movie files that were approximately 45 seconds long. We loaded all stimulus movie files into PsychoPy and E-Prime using the instructions on OSF.

Procedure. We used E-Prime to present stimulus movie files for data collection efforts reported here; PsychoPy could also be used. Participants completed the facial expression-decoding task as well as a mental rotation task (Peters, 1995; data not reported here) in a randomized order. The facial expression-decoding task requires participants to view 34 degraded, pixelated images of 17 male and 17 female faces displaying one of seven emotions (happy, angry, surprise, fear, sad, disgust, neutral) on a computer monitor connected to a keyboard for response input. In each trial, a single face, randomly selected from the set of 34 faces, slowly comes into focus over the course of approximately 45 seconds. Participants were required to halt the trial by pressing the space bar as soon as they recognize the displayed emotion. A face mask was then presented for 500 milliseconds. A screen with a list of seven emotion response options was automatically displayed after the mask. The trial ended when participants indicated which one of seven emotion options they saw by pressing the numeric key (1-7) that corresponded to the number next to the emotion listed. The next trial began immediately after a number was entered by the participant. Primary dependent variables in the task include facial-expression recognition accuracy and reaction time. Figure 2 shows an example stimulus gradually coming into focus using static frames. Table 3 depicts a breakdown of the number of male and female faces used in the task, as well as the emotion displayed by each face.

Data cleaning and analytic approach. Data were screened on a trial-level basis. Out of 14,112 trials across subjects, 629 trials were excluded from implausibly low reaction times (less than 500 *ms*; i.e., they responded before any pixelated noise had been removed from the degraded image), and 238 outliers were excluded (exceeding three standard deviations of the scaled reaction time data). The final sample included 13,245 trials across all subjects. No transformations were applied to the reaction time data (Schramm & Rouder, 2019). Using the *lme4* package in R (Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2018), trial-level mixed effects models were estimated to test sex differences in both reaction times and accuracy. We also tested the interaction between participant sex and emotion displayed in the stimuli, and participant race. Models were estimated using random participant intercepts.

Results

As shown in Figure 3, women were more accurate (70% correct) than men (65% correct) in facial emotion decoding, $t(399.87) = 4.08$, $p < 0.001$, $d = 0.41$. Women (2210.70 *ms*) also responded more quickly than did men (2296.32 *ms*), $t(422.14) = 2.08$, $p = 0.04$, $d = 0.20$. The female advantage in reaction times was largely due to faster processing of male ($p = 0.02$, $d = 0.23$) than female ($p = 0.14$, $d = 0.15$) faces. In contrast, the female advantage in accuracy was consistent across male ($p < 0.001$, $d = 0.38$) and female ($p < 0.001$, $d = 0.34$) faces.

We observed no interaction effect between participant sex and emotion displayed in the stimuli, $F(2,12855.7) = 1.21$, $p = 0.296$, or participant race, $F(2,12985) = 0.11$, $p = 0.892$, on response times. Likewise, there was no interaction effect between sex and participant race on accuracy, $F(2,13237) = 0.42$, $p = .655$. There was, however, a significant interaction effect between participant sex and emotion displayed in the stimuli on accuracy, $F(6,12082.7) = 7.15$, $p < 0.001$. Table 4 lists sex differences across the seven emotions displayed by stimuli. This interaction indicated that the sex differences in accuracy were mainly driven by the disgust, fear, neutral, and sad emotions.

Conclusion

Overall, this example demonstrates the utility of this type of facial expression-decoding task using a degraded image paradigm. This task revealed moderate sex differences that are consistent in direction and magnitude to those found in previous studies (J. K. Hall et al., 2010; Hoffmann et al., 2010; Sasson et al., 2010) that in turn supports the external validity of the current assessment and procedure. The findings support the use of this novel task for studying sex differences in social acuity.

Discussion

Although there are many published versions of facial expression-decoding tasks, and degraded image tasks broadly, they are limited to use of a single face database or a database that that cannot be readily tailored to test specific experimental hypotheses. Moreover, many traditional social acuity tasks use morphed faces that may seem unnatural or can be costly to construct. To circumvent these issues, a user may opt to create a new facial expression-decoding task or degraded image paradigm, but this can be effortful and time-consuming. The aim of the present tutorial was to provide a step-by-step, user-friendly guide to creating a degraded image task using freely available resources. We provide interactive resources to create experiments adaptable for many research questions, including an example facial expression-decoding task. The goal was to allow users to quickly and effortlessly develop, implement, and analyze future degraded image experiments and answer new research questions without the delay of time between conceptualization of a new task and evaluation of experimental data.

Limitations

Unlike many existing face emotion recognition and perception tasks, the facial expression-decoding task described here cannot be used as a diagnostic neuropsychological test because it is not standardized. Accordingly, the data from our example cannot be used as normative data. The user should also be aware that sex differences in color perception and visual acuity could bias sex differences in facial expression-decoding tasks using noisy stimuli because color plays a role in perception when shapes are degraded (Sinha, Balas, Ostrovsky, & Russell, 2006). Specifically, of those with Northern European ancestry (like those who comprise our sample), approximately 8% of men and 0.5% percent of women have red-green color blindness (National Eye Institute, 2015). Given only a small percentage of men in our sample were potentially colorblind, and with effect sizes falling within the rage of sex differences in facial expression-decoding, the inclusion of colorblind men likely did not substantially alter the signal-to-noise ratio in the data, but we caution the user to consider excluding those with red-green color blindness if color images are used in a degraded image task.

We note that Grady et al. (1996), Royer, Blais, Gosselin, Duncan, and Fiset (2015), and Sadr & Sinha (2004) used black and white degraded face images whereas Rossion and Caharel (2011), Rossion, Hanseeuw, and Dricot (2012), and Wentura and Rohr (2018) used color face images in their degraded image research, and refer the user to this prior research when deciding whether to use color or black and white images, depending on the goal of the task. Users who degrade images other that faces may also benefit from reviewing prior research on degraded objects and words. Finally, we note that we did not screen for normal or corrected-to-normal vision, which could potentially confound in the findings reported here.

Despite these limitations, we found sex differences in speed and accuracy that are consistent with previous studies (J. K. Hall et al., 2010; Hoffmann et al., 2010; Sasson et al., 2010), indicating that the task is externally valid and useful for assessing group differences in face processing. The tutorial provides a user with the tools to create and carry out a wide array of experiments, and the flexibility to adapt this task in new ways at no cost to the user. Additionally, new methods are available that allow researchers to construct tasks using stimuli from large data sets while providing guidance on how to select stimuli to improve validity (Hsu, Martin, Sanborn, & Griffiths, 2019).

Conclusion

Perceptual identification tasks are applicable broadly to the study of psychiatric and neurological disorders (Johnston, Stojanov, Devir, & Schall, 2005) and to tracking developmental changes throughout childhood and adulthood (Herba & Phillips, 2004). They are also of interest to evolutionary-minded scientists (Geary, 2015) and relevant in areas of the social and cognitive neurosciences (Seitz et al., 2008), including how face processing is specialized (Wong, Palmeri, & Gauthier, 2009). We believe this detailed tutorial, including access to our current task on multiple platforms, as well as instructions for editing the task and extracting data, will be of use to a broad range of investigators. For example, we hope this tutorial will be especially useful to new investigators who otherwise might have difficulty with creating this type of task. We also anticipate this tutorial will be of interest to the broad audience of psychologists and others in search of degraded image (e.g., social acuity) tasks. Among its many applications, this task can be used to test implicit bias, stereotypes, visual processing, evolutionary theory, and cognitive vulnerabilities. We hope our tutorial will encourage and facilitate future research on many topics.

Transparency and Openness

All programmed tasks, instructional content, code, and data can be found on both OSF (https://osf.io/ntmah/) and Github (https://github.com/scofieldjohn/degraded_image_tutorial).

References

Adolphs, R., Baron-Cohen, S., & Tranel, D. (2002). Impaired recognition of social emotions following amygdala damage. *Journal of Cognitive Neuroscience*, *14*(8), 1264–1274.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Letters To Nature*, *372*, 669–672. https://doi.org/10.1038/nature02251.1.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Brown, D. E. (1991). *Human Universals*. New York City: McGraw-Hill.

Buck, R. W., Savin, V. J., Miller, R. E., & Caul, W. F. (1972). Communication of affect through facial expressions in humans. *Journal of Personality and Social Psychology*, *23*(3), 362–371. https://doi.org/10.1037/h0033171

Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal decoding skills and relationship well-being in adults. *Journal of Nonverbal Behavior*, *23*(1), 91–100. https://doi.org/10.1023/A:1021339410262

Darwin, C. (1872). *The expression of emotions in man and animals*. London: William Clowes & Sons. https://doi.org/10.1037/10025-007

Del Zotto, M., & Pegna, A. J. (2015). Processing of masked and unmasked emotional faces under different attentional conditions: An electrophysiological investigation. *Frontiers in Psychology*, *6*(OCT), 1–12. https://doi.org/10.3389/fpsyg.2015.01691

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S. J., & Friston, K. J. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature*, *389*(6651), 596–599. https://doi.org/10.1038/39309

Domes, G., Schulze, L., & Herpertz, S. C. (2009). Emotion recognition in borderline personality disorder: A review of the literature. *Journal of Personality Disorders*, *23*(1), 6–19.

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing black: Race, crime, and visual processing. *Journal of Personality and Social Psychology*, *87*(6), 876–893. https://doi.org/10.1037/0022-3514.87.6.876

Edwards, J., Jackson, H. J., & Pattison, P. E. (2002). Emotion recognition via facial expression and affective prosody in schizophrenia: A methodological review. *Clinical Psychology Review*, *22*(8), 789–832.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin*, *115*, 268–287.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–129. https://doi.org/10.1037/h0030377

Elfenbein, H. A., Marsh, A. A., & Ambady, N. (2002). Emotional intelligence and the recognition of emotion from facial expressions. *The Wisdom in Feeling: Psychological Processes in Emotional Intelligence Emotions and Social Behavior*, 37–59.

Eslinger, P. J., & Damasio, A. R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. *Neurology*, *35*(12), 1731–1741.

Fazio, R. H., Williams, C. J., & Powell, M. C. (2000). Associative Strength: Measuring Activation from and Their Associations Memory. *Political Psychology*, *21*(1), 7–25.

Geary, D. C. (2015). *Evolution of vulnerability: Implications for sex differences in health and development*. Academic Press.

Geary, D. C., Winegard, B., & Winegard, B. (2014). Reflections on the evolution of human sex differences: Social selection and the evolution of competition among women. In *olutionary perspectives on human sexual psychology and behavior* (pp. 393–412). New York, NY: Springer.

Ghiselin, M., Ekman, P., & Gruber, H. E. (1974). Darwin and Facial Expression: A Century of Research in Review. *Systematic Zoology*. https://doi.org/10.2307/2412481

Gollin, E. S. (1960). Developmental studies of visual recognition of incomplete objects. *Percept*, *11*, 289–298.

Grady, C. L., Horwitz, B., Pietrini, P., Mentis, M. J., Ungerleider, L. G., Rapoport, S., & Haxby, J. V. (1996). Effect of task difficulty on cerebral blood flow during perceptual matching of faces. *Human Brain Mapping*, *4074*, 227–239.

Hall, J. A., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion*, *4*(2), 201–206. https://doi.org/10.1037/1528-3542.4.2.201

Hall, J. K., Hutton, S. B., & Morgan, M. J. (2010). Sex differences in scanning faces: Does attention to the eyes explain female superiority in facial expression recognition? *Cognition and Emotion*, *24*(4), 629–637. https://doi.org/10.1080/02699930902906882

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, *20*(3), 290–322. https://doi.org/10.1007/s11065-010-9138-6

Herba, C., & Phillips, M. (2004). Annotation: Development of facial expression recognition from childhood to adolescence: Behavioural and neurological perspectives. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *45*(7), 1185–1198. https://doi.org/10.1111/j.1469-7610.2004.00316.x

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior*, *21*(4), 241–257.

Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H. C. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta Psychologica*, *135*(3), 278–283. https://doi.org/10.1016/j.actpsy.2010.07.012

Hsu, A. S., Martin, J. B., Sanborn, A. N., & Griffiths, T. L. (2019). Identifying category representations for complex stimuli using discrete Markov chain Monte Carlo with people. *Behavior Research Methods*. https://doi.org/10.3758/s13428-019-01201-9

Izard, C. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.

Izard, C. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin*, *115*, 288–299.

James, T. W., Goodale, M. A., Menon, R. S., Humphrey, G. K., & Gati, J. S. (2000). The effects of visual object priming on brain activation before and after recognition. *Current Biology*, *10*, 1017–1024. https://doi.org/10.1016/s0960-9822(00)00655-2

Johnston, P. J., Stojanov, W., Devir, H., & Schall, U. (2005). Functional MRI of facial emotion recognition deficits in schizophrenia and their electrophysiological correlates. *European Journal of Neuroscience*, *22*(5), 1221–1232. https://doi.org/10.1111/j.1460-9568.2005.04294.x

Kessler, H., Bayerl, P., Deighton, R. M., & Traue, H. C. (2002). Facially Expressed Emotion Labeling (FEEL): A computer-based test for emotion recognition Summary. *Verhaltenstherapie & Verhaltensmedizin*, (3), 297–306.

Koster, E. H. W., Verschuere, B., Burssens, B., Custers, R., & Crombez, G. (2007). Attention for emotional faces under restricted awareness revisited: Do emotional faces automatically attract attention? *Emotion*, *7*(2), 285–295. https://doi.org/10.1037/1528-3542.7.2.285

Kret, M. E., & De Gelder, B. (2012). A review on sex differences in processing emotional signals. *Neuropsychologia*, *50*, 1211–1221. https://doi.org/10.1016/j.neuropsychologia.2011.12.022

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud faces database. *Cognition and Emotion*, *24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces (KDEF). *Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet*.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*, 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Macrae, C. N., Stangor, C., & Milne, A. B. (1994). Activating social stereotypes: A functional analysis. *Journal of Experimental Social Psychology*, *30*, 370–389.

Mahr, T. (2015). Package ‘rprime’: Functions for working with ‘Eprime’ text files. Retrieved from https://cran.r-project.org/package=rprime

Matsuda, Y.-T., Fujimura, T., Katahira, K., Okada, M., Ueno, K., Cheng, K., & Okanoya, K. (2013). The implicit processing of categorical and dimensional strategies: An fMRI study of facial emotion perception. *Frontiers in Human Neuroscience*, *7*, 1–20. https://doi.org/10.3389/fnhum.2013.00551

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin*, *126*(3), 424–453.

McDermott, R. (2011). Internal and external validity. In J. N. Druckman, D. P. Green, J. H. Kuklinski, & A. Lupia (Eds.), *Cambridge Handbook of Experimental Political Science* (pp. 27–40). New York: Cambridge University Press.

National Eye Institute. (2015). Facts About Color Blindness. Retrieved from https://nei.nih.gov/health/color_blindness/facts_about

Oldershaw, A., Hambrook, D., Stahl, D., Tchanturia, K., Treasure, J., & Schmidt, U. (2011). The socio-emotional processing stream in Anorexia Nervosa. *Neuroscience and Biobehavioral Reviews*, *35*(3), 970–988. https://doi.org/10.1016/j.neubiorev.2010.11.001

Pantic, M., Valstar, M., Rademaker, R., & Maat, L. (2005). Web-based database for facial expression analysis. In *Proc. IEEE Int'l Conf. on Multimedia and Expo (ICME'05)*. Amsterdam, The Netherlands. https://doi.org/10.1109/ICME.2005.1521424

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 3,* 1–9. https://doi.org/10.3758/s13428-018-01193-y

Peters, M. (1995). Mental Rotations Test (MRT-A).

Pollatos, O., Herbert, B. M., Schandry, R., & Gramann, K. (2008). Impaired central processing of emotional faces in Anorexia Nervosa. *Psychosomatic Medicine*, *70*(6), 701–708. https://doi.org/10.1097/PSY.0b013e31817e41e6

Psychology Software Tools. (2015). *Chronos Operator Manual*. Pittsburgh, PA: Psychology Software Tools, Inc.

Psychology Software Tools. (2016). *E-Prime 3.0*. Pittsburgh, PA: Psychology Software Tools, Inc. Retrieved from www.pstnet.com

R Core Team. (2018). *R: A language and environment for statistical computing*. *R Foundation* *for Statistical Computing*. Vienna, Austria. Retrieved from https://www.r-project.org/

Richard, L., & Charbonneau, D. (2009). An introduction to E-Prime. *Tutorials in Quantitative Methods for Psychology*, *5*(2), 68–76. https://doi.org/10.20982/tqmp.05.2.p068

Rosenthal, R., Archer, D., Hall, J. A., DiMatteo, M. R., & Rogers, P. L. (1979). Measuring sensitivity to nonverbal communication: The PONS test. In *Nonverbal Behavior* (pp. 67–98). Academic Press.

Rossion, B. (2014). Understanding face perception by means of prosopagnosia and neuroimaging. *Frontiers in Bioscience*, *6*, 258–307.

Rossion, B., & Caharel, S. (2011). ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception. *Vision Research*, *51*(12), 1297–1311. https://doi.org/10.1016/j.visres.2011.04.003

Rossion, B., Hanseeuw, B., & Dricot, L. (2012). Defining face perception areas in the human brain: A large-scale factorial fMRI face localizer analysis. *Brain and Cognition*, *79*(2), 138–157. https://doi.org/10.1016/j.bandc.2012.01.001

Royer, J., Blais, C., Gosselin, F., Duncan, J., & Fiset, D. (2015). When less is more: Impact of face processing ability on recognition of visually degraded faces. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(5), 1179–1183. https://doi.org/10.1037/xhp0000095

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience and Biobehavioral Reviews*, *32*(4), 863–881. https://doi.org/10.1016/j.neubiorev.2008.01.001

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, *115*(1), 102–141. https://doi.org/10.1037/0033-2909.115.1.102

Sadr, J., & Sinha, P. (2004). Object recognition and Random Image Structure Evolution. *Cognitive Science*, *28*(2), 259–287. https://doi.org/10.1016/j.cogsci.2003.09.003

Sasson, N. J., Pinkham, A. E., Richard, J., Hughett, P., Gur, R. E., & Gur, R. C. (2010). Controlling for Response Biases Clarifies Sex and Age Differences in Facial Affect Recognition. *Journal of Nonverbal Behavior, 34*(4), 207–221. https://doi.org/10.1007/s10919-010-0092-z

Schramm, P., & Rouder, J. (2019). Are Reaction Time Transformations Really Beneficial. Retrieved from https://doi.org/10.31234/osf.io

Schultz, J., & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research*, *194*, 465–475. https://doi.org/10.1007/s00221-009-1721-9

Seitz, R. J., Schäfer, R., Scherfeld, D., Friederichs, S., Popp, K., Wittsack, H. J., … Franz, M. (2008). Valuating other people's emotional face expression: A combined functional magnetic resonance imaging and electroencephalography study. *Neuroscience*, *152*(3), 713–722. https://doi.org/10.1016/j.neuroscience.2007.10.066

Sinha, P., Balas, B., Ostrovsky, Y., & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE*, *94*(11), 1948–1961. https://doi.org/10.1109/JPROC.2006.884093

Snodgrass, J. G., Smith, B., Feenan, K., & Corwin, J. (1987). Fragmenting pictures on the apple macintosh computer for experimental and clinical applications. *Behavior Research Methods, Instruments, & Computers*, *19*(2), 270–274. https://doi.org/10.3758/BF03203798

Stone, V. E., Baron-Cohen, S., Calder, A., Keane, J., & Young, A. (2003). Acquired theory of mind impairments in individuals with bilateral amygdala lesions. *Neuropsychologia*, *41*(2), 209–220. https://doi.org/10.1016/S0028-3932(02)00151-3

Strohminger, N., Gray, K., Chituc, V., Heffner, J., Schein, C., & Heagins, T. B. (2015). The MR2: A multi-racial, mega-resolution database of facial stimuli. *Behavior Research Methods*, *48*, 1197–1204. https://doi.org/10.3758/s13428-015-0641-9

The GIMP Team. (2019a). GNU Image Manipulation (GIMP) Version 2.10. Retrieved from www.gimp.org

The GIMP Team. (2019b). GNU Image Manipulation (GIMP) Version 2.10.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., … Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*, 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*(1), 35–57.

Valstar, M. F., & Pantic, M. (2010). Induced disgust, happiness and surprise: An addition to the MMI Facial Expression Database. In *Proceedings of the International Language Resources and Evaluation Conference* (pp. 65–70). Malta.

van Beek, Y., & Dubas, J. S. (2008). Age and gender differences in decoding basic and non-basic facial expressions in late childhood and early adolescence. *Journal of Nonverbal Behavior*, *32*, 37–52. https://doi.org/10.1007/s10919-007-0040-8

Wager, T. D., Phan, K. L., Liberzon, I., & Taylor, S. F. (2003). Valence, gender, and lateralization of functional brain anatomy in emotion: A meta-analysis of findings from neuroimaging. *NeuroImage*, *19*(3), 513–531. https://doi.org/10.1016/S1053-

8119(03)00078-8

Wentura, D., & Rohr, M. (2018). Emotion-specific priming effects with marginally perceptible facial expression primes: Evidence from the “leave-one-out” paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(12), 1946–1969. https://doi.org/10.1037/xhp0000581

Wilhelm, O., Hildebrandt, A., Manske, K., Schacht, A., & Sommer, W. (2014). Test battery for measuring the perception and recognition of facial expressions of emotion. *Frontiers in Psychology*, *5*, 1–23. https://doi.org/10.3389/fpsyg.2014.00404

Wong, A. C.-N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects. *Psychological Science*, *20*(9), 1108–1117. https://doi.org/10.1111/j.1467-9280.2009.02430.x

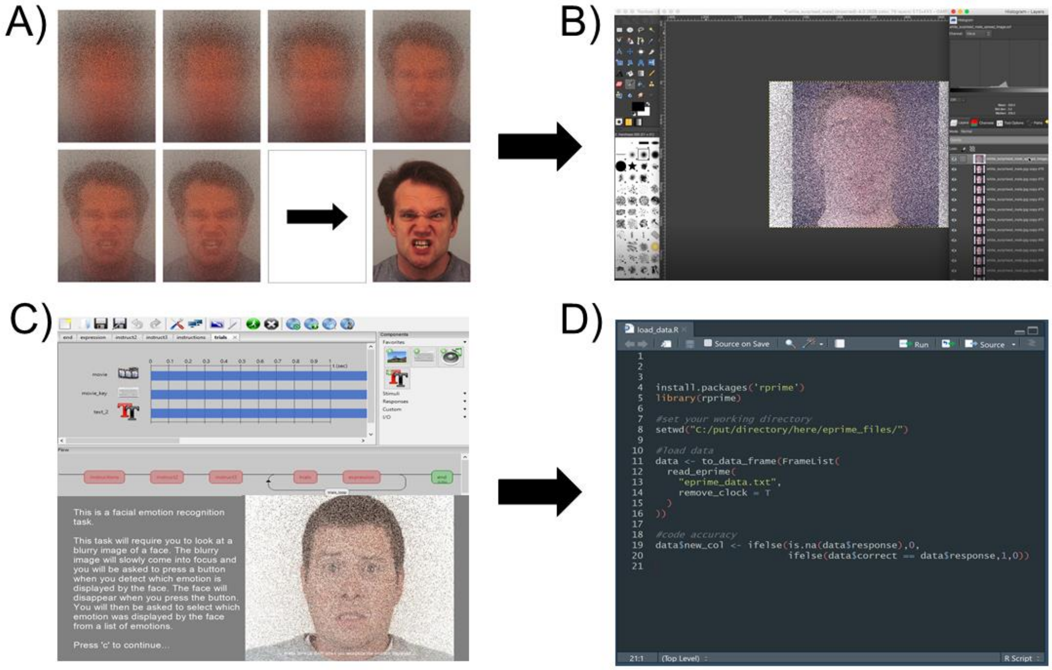

Figure 1. Tutorial overview: A) Degrade images by adding pixelated noise; B) Create stimulus movie files; C) Use PsychoPy or E-Prime to Present stimuli for data collection; D) Extract, convert, and aggregate data using *R*.

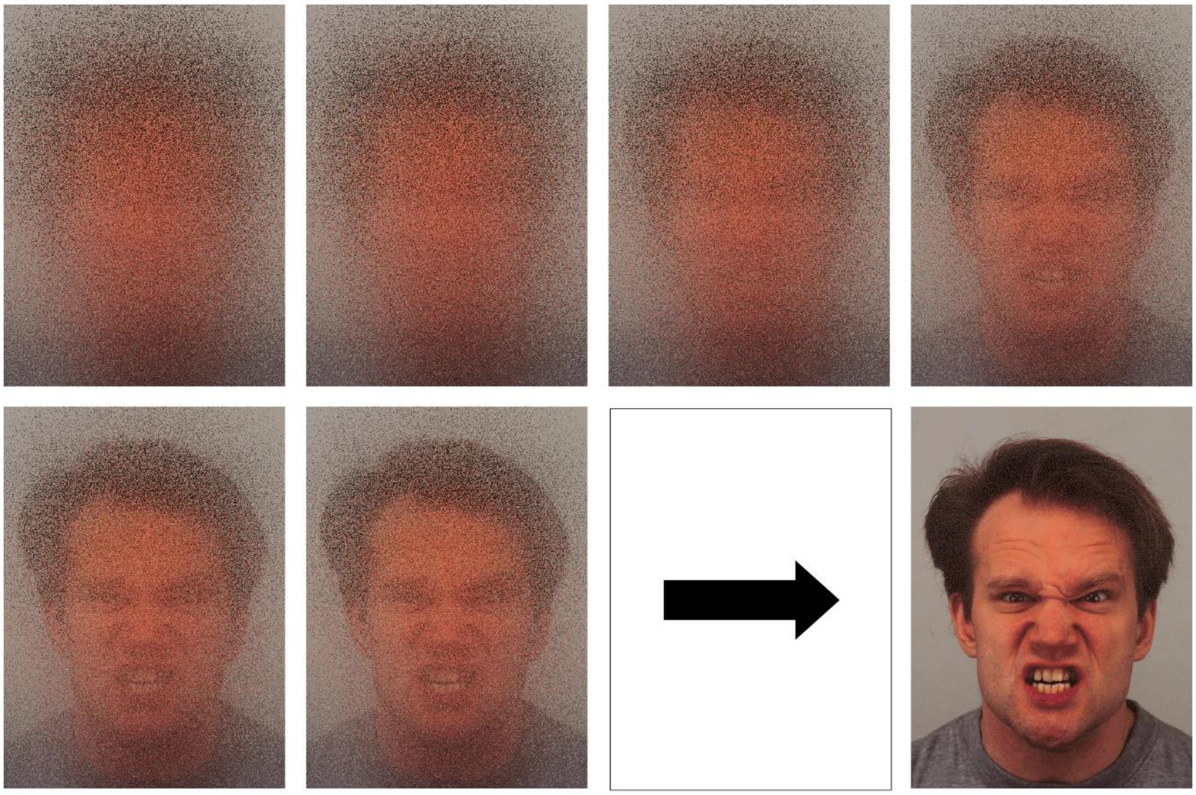

Figure 2. Static screenshots from an example degraded image stimulus movie file displaying a degraded image of an angry male face. The stimulus progresses through 78 frames across approximately 45 seconds, beginning with the example frame in the top left and ending with the example frame in the bottom right.

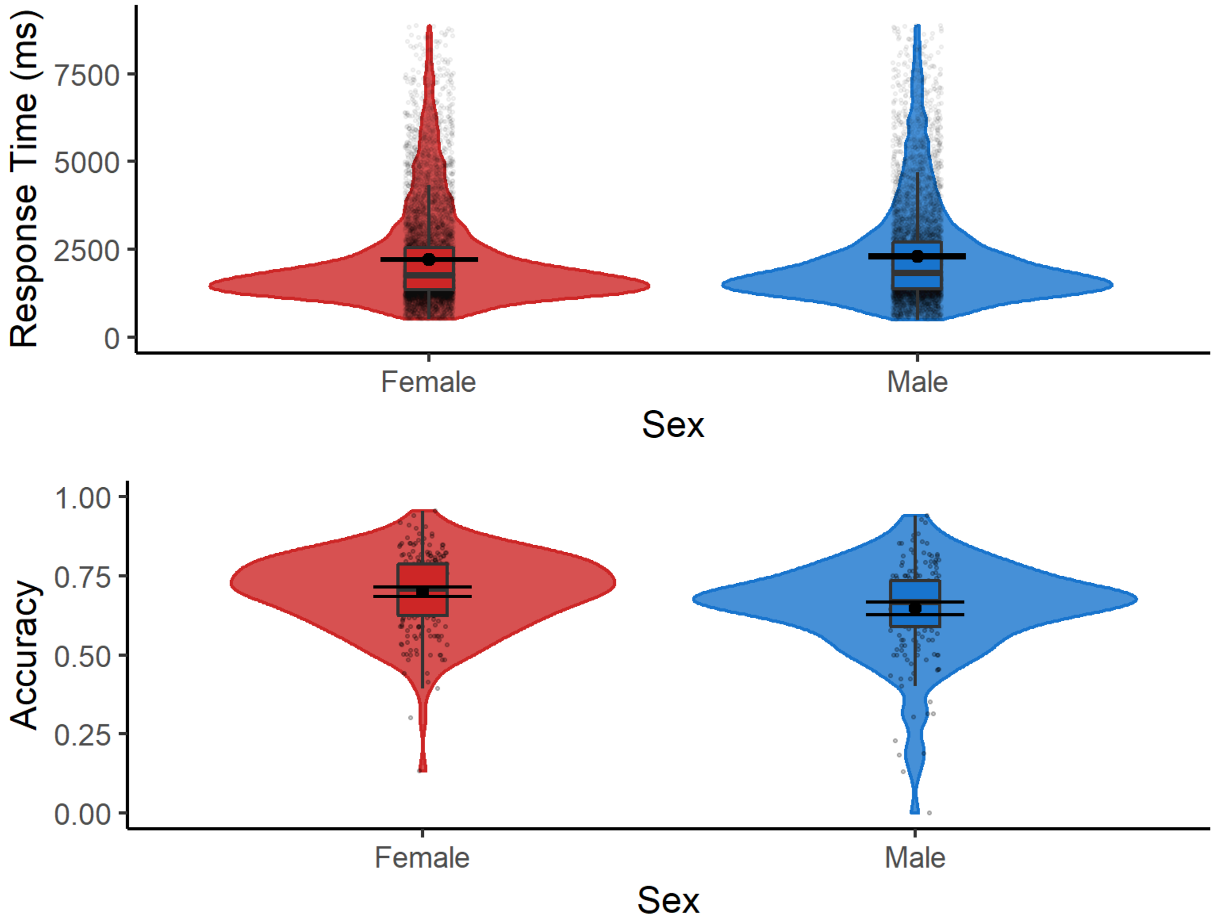

Figure 3. Violin plots showing sex differences in response times (top panel) and accuracy (bottom panel) in the facial-expression decoding task.

Table 1. Face databases used in example.

Face Database	Reference	URL	Description
MMI Facial Expression Database	Pantic, M., Valstar, M., Rademaker, R., & Maat, L. (2005). Web-based database for facial expression analysis. In *Proc. IEEE Int'l Conf. on Multimedia and Expo (ICME'05)*. Amsterdam, The Netherlands. Valstar, M. F., & Pantic, M. (2010). Induced disgust, happiness and surprise: An addition to the MMI Facial Expression Database. In *Proceedings of the International Language Resources and Evaluation Conference* (pp. 65–70). Malta.	https://mmifacedb.eu	2,900 high-resolution still images of 75 individuals. Six basic emotions.
Karolinska Directed Emotional	Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional	http://kdef.se	4900 pictures of 70 individuals.

Faces (KDEF)	Faces (KDEF). *Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet*.		7 emotional expressions.
Radboud Faces Database (RaFD)	Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud faces database. *Cognition and Emotion*, *24*(8), 1377–1388.	http://rafd.nl	67 Caucasian adult and juvenile males and females, and Moroccan Dutch males. 8 emotional expressions.
Chicago Face Database	Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*, 1122–1135.	https://chicagofaces.org	158 Black and White males and females between the ages of 17-65. Neutral, happy, threatening, and fearful expressions.

Table 2. Software tools.

Software	URL	Purpose

GIMP v.2.10.8	https://gimp.org	Image Editor
PsychoPy3 v.3.1.5	https://psychopy.org	Stimulus Presentation
E-Prime3	https://pstnet.com/products/e-prime	Stimulus Presentation
R v.3.5.1	https://r-project.org	Data Analysis

Table 3. Characteristics of 34 faces used in degraded image stimulus movie files, including 17 male and 17 female faces, with at least two male faces and at least two female faces each displaying one of seven emotions.

Image Number	Emotion	Sex
1	Angry	Female
2	Angry	Female
3	Angry	Male
4	Angry	Male
5	Disgust	Female
6	Disgust	Female
7	Disgust	Female
8	Disgust	Male
9	Disgust	Male
10	Disgust	Male
11	Fear	Female
12	Fear	Female
13	Fear	Male
14	Fear	Male
15	Happy	Female
16	Happy	Female
17	Happy	Female
18	Happy	Male
19	Happy	Male

20	Happy	Male
21	Neutral	Female
22	Neutral	Female
23	Neutral	Female
24	Neutral	Male
25	Neutral	Male
26	Neutral	Male
27	Sad	Female
28	Sad	Female
29	Sad	Male
30	Sad	Male
31	Surprise	Female
32	Surprise	Female
33	Surprise	Male
34	Surprise	Male

Table 4. Interaction effects.

Emotion	*p* value	Cohen's d
Anger	0.350	-0.098
Disgust	0.002	0.353
Fear	0.011	0.315
Happy	0.999	0.038
Neutral	< 0.001	0.477
Sad	0.041	0.274
Surprise	0.999	0.038

Note. This table shows sex differences for various emotions. *p* values are bonferroni corrected to control for multiple comparisons. Negative Cohen’s *d* values indicate accuracy is higher in males versus females, and positive *d* values indicate that accuracy is higher in females than males.