Introducing the Fear Learning and Anxiety Response (FLARe) app and web portal for the remote delivery of fear conditioning experiments

*T. McGregor¹, MSc., *K. L. Purves^{1,2}, PhD., T. Barry³, PhD., E. Constantinou^{1,4}, PhD., M. G. Craske⁵, PhD., G. Breen^{1,2}, PhD., ^K. S. Young^{1,2}, DPhil., ^T. C. Eley^{1,2}, PhD.

¹King's College London; Social, Genetic and Developmental Psychiatry Centre; Institute of Psychiatry, Psychology & Neuroscience; London, UK.

²NIHR Biomedical Research Centre for Mental Health; South London and Maudsley NHS Trust; London, UK.

³Experimental Psychopathology Lab; Department of Psychology; The University of Hong Kong, Hong Kong.

⁴Department of Psychology, University of Cyprus; Nicosia, Cyprus.

⁵Department of Psychology and Department of Psychiatry and Biobehavioral Sciences; University of California, Los Angeles; California, USA.

*Authors contributed equally

^Authors contributed equally

Code availability

<https://github.com/flare-kcl>

Abstract

Experimental paradigms measuring key psychological constructs can enhance our understanding of mechanisms underlying human psychological well-being and mental health. Delivering such paradigms remotely affords opportunities to reach larger, more representative samples than is typically possible with in-person research. The efficiency gained from remote delivery makes it easier to test replication of previously established effects in well-powered samples. There are several challenges to the successful development and delivery of remote experimental paradigms, including use of an appropriate delivery platform, identifying feasible outcome measures, and metrics of participant compliance. In this paper, we present FLARe (Fear Learning and Anxiety Response), open-source software in the form of a smartphone app and web portal for the creation and delivery of remote fear conditioning experiments. We describe the particular benefits and challenges associated with the creation of a remote delivery platform for fear conditioning, before presenting in detail the resultant software suite, and one particular instance of deploying this using the FLARe research infrastructure. We provide examples of the application of FLARe to a number of research questions which illustrate the particular benefits of the remote approach to experiment delivery. The FLARe smartphone app and web portal are available for use by other researchers and have been designed to be user friendly and intuitive. We hope that FLARe will be a useful tool for those interested in conducting well-powered fear conditioning studies to inform our understanding of the development and treatment of anxiety disorders.

Introduction

Background

Experimental paradigms are a fundamental methodology within psychological science. A carefully controlled experiment that measures how individuals react to specific stimuli can uncover the mechanisms that underlie psychological constructs and help establish cause and effect.

Laboratory-based experimental testing, however, is constrained by the significant time, financial costs and manpower needed to conduct in-person assessments with individual participants. These constraints often limit the sample sizes of experimental research, resulting in many studies being underpowered. Underpowered research reduces the chance of detecting true effects (more ‘false negatives’), and increases the chance of “significant” effects occurring by chance alone (more ‘false positives’; Ioannidis, 2005; Munafò et al., 2017; Open Science Collaboration, 2015). In-person research can face particular challenges in achieving representative samples. Researchers are restricted to individuals who are nearby or can afford to travel or to easily accessible groups like psychology students. Together, these limitations harm the generalisability and replicability of experiments conducted in-person.

Remote experiments affords opportunities to collect much larger samples in shorter periods of time, with participants completing tasks on their own personal computers and mobile devices, with no experimenter interaction (Brown et al., 2014; de Leeuw, 2015; Gureckis et al., 2016; Jardine, Fisher, & Carrick, 2015; Sochat et al., 2016). Well-designed digital research tools can also reduce the burden on researchers by making it easier to create, share, and configure psychological experiments. Furthermore, the software can be released with source code (open-source software) to help increase access, enable replication, reduce duplication of effort, and encourage innovation and expansion. Online, open-source research tools can also serve to democratise research, allowing research groups who cannot afford expensive specialist equipment and software to conduct experiments that help them answer their research questions in well-powered samples (Adjerid & Kelley, 2018).

Development of experimental research tools to date have taken two different approaches, either non-specific, or task-specific. Non-specific tools are highly adaptable, allowing researchers to create a range of tasks on a single platform (e.g. Gorilla; Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020). Task-specific tools are created with a specific paradigm in mind and only allow flexibility that is relevant to the task whilst implementing control over important experimental

parameters (e.g. SMARTA, RECAPXPR; Cortis Mack, Harding, Davies, & Ward, 2019; Vagell, Vagell, Jacobs, Gordon, & Baden, 2019).

A key strength of non-specific tools is their flexibility, which allows adaptation of almost any experimental task into an easily deliverable online format. However, achieving this flexibility often requires a sacrifice in accessibility. For example, non-specific research tools may require the user to build their experiment using the command line, or without any paradigm-specific guidance on the parameters that need to be considered. Furthermore, data derived from non-specific research tools may require substantial processing to extract the parameters of interest within a particular paradigm. As such, adapting a task for remote research using a non-specific research tool can be difficult for those who are new to coding or new to the task itself, requiring additional user expertise and time. Additionally, non-specific tools may not allow for strict control or measurement of the experimental environment that can facilitate standardisation of experimental tasks outside of the laboratory (e.g. volume or brightness settings on a smartphone).

Task-specific research tools can be developed in a user-friendly way that does not require coding expertise, improving accessibility for researchers and facilitating standardisation of methods between in-person and online delivery. Whilst this approach inherently places more constraints on the user than non-specific research tools (e.g. Gorilla; Anwyl-Irvine et al., 2020), it presents a compromise between flexibility and usability. Maintaining limits on the range of parameters that can be altered across studies ultimately encourages greater methodological consistency.

Here we describe the Fear Learning and Anxiety Response (FLARe) smartphone app and associated web portal, a set of task-specific tools designed for conducting fear conditioning research remotely. A key priority of the FLARe project was to provide an open-source, user-friendly platform to provide access to a validated remote fear conditioning measure. As such, we opted for a task-specific approach. Overall, creating a task-specific, semi-flexible, remote research tool has the advantage that many of the technological and scientific issues involved in setting up, programming and managing tasks are addressed by the developers. This approach streamlines the process of study design, implementation and data acquisition for other researchers.

Fear conditioning

Fear conditioning is a widely used experimental paradigm that models the development, maintenance, treatment, and relapse of anxiety disorders (Craske, Hermans, & Vervliet, 2018;

Mineka & Oehlberg, 2008). Meta-analyses and large-scale experiments have found reliable associations with anxiety disorders (Duits et al., 2015; Lissek et al., 2005; McGregor et al., 2021) that replicate across cohorts. Basic fear conditioning paradigms consist of two phases known as ‘acquisition’ and ‘extinction’. Acquisition models the development of fears via direct learning experiences, whereas extinction models the subsequent learning of safety and is considered to be a central mechanism of exposure-based treatments.

The paradigm relies on principles of Pavlovian conditioning in order to create associations between different stimuli (**Figure 1**). In **acquisition**, participants are presented with neutral ‘conditional stimuli’ (CS; e.g. images of different sized circles). One CS is repeatedly paired with an aversive ‘unconditional stimulus’ (US; e.g. an unpleasant loud noise), while the other is always presented alone. Through repeated trials, participants learn to associate one of the CSs with threat due to its pairing with the US. This CS is known as the ‘CS+’. The other CS, not paired with the US, becomes associated with safety and is known as the ‘CS-’. During the **extinction** phase, both CSs are repeatedly presented in the absence of the US. Extinction leads to new learning, with the CS+ now holding competing meanings of threat and safety (Bouton, 1993). Throughout each stage, a participant’s ability to inhibit their fear response to a safety cue (CS-) relative to a threatening cue (CS+) gives an index of ‘discriminant conditioning’.

Two additional phases can be used in combination with acquisition and extinction to explore other anxiety-related processes: generalisation and renewal. The **generalisation** phase models how fear learning transfers to stimuli with similar perceptual or conceptual properties to the CS+ and CS- (Dymond, Dunsmoor, Vervliet, Roche, & Hermans, 2015). Generalisation of fear learning to different sized circles could be assessed by presenting multiple circles that range in size between the circles used for the CS+ and CS-. Generalisation is assessed immediately after acquisition, prior to extinction. **Renewal** models the return of fear after treatment and is assessed typically at least 24 hours after acquisition and extinction take place (Lonsdorf et al., 2017). During the renewal phase participants are presented with the CS+ and CS- again, without the US (similarly to the extinction phase). This phase tests the extent to which participants retrieve their extinction learning over an extended period of time.

Figure 1. The stages of fear conditioning. CS = conditional stimulus, GS = generalisation stimulus, US = unconditional stimulus.

Fear conditioning can be quantified in a number of ways by measuring physiological, behavioural, self-report, and neurobiological responses. Specifically, skin conductance responses (physiological), avoidance (behavioural), expectancy ratings (self-report), and functional neuroimaging (neurobiological) are typical outcome measures in the field of fear conditioning (for more detail see Lonsdorf et al., 2017).

FLARe rationale

The FLARe software was developed by a team based primarily at King's College London to remotely deliver a fear conditioning task. Our primary research aim was to facilitate large-scale data collection necessary for evaluating genetic and environmental influences upon fear conditioning. Given the potential for 'big-data' to improve reliability of psychological research (Adjerid & Kelley, 2018), and the pivotal role of open and reproducible scientific practices, a second key aim was to create a platform available to other researchers without the necessity for advanced programming expertise or high budgets.

Choice of delivery platform

Remote tasks can be completed by participants on computers, tablets or smartphones. Though many platforms exist to support the development of computer-based tasks (e.g. Anwyl-Irvine et al., 2020; de Leeuw, 2015; Gureckis et al., 2016; Lange, Kühn, & Filevich, 2015; Sochat et al., 2016) the proportion of the population that owns a smartphone is increasing and has overtaken laptop/desktop ownership in some countries (e.g. the United States Pew Research Center, 2021). Additionally, smartphones have a range of available functionality (e.g. passive data monitoring of app state, phone volume and in-built notification flags) that can be used for a priori and post hoc control of the remote testing environment. For these reasons, we developed a smartphone app to enable remote experiment delivery to participants and a partner web-platform (portal) for experiment setup and data processing for researchers.

Choice of outcome measures

We selected self-reported fear conditioning outcomes, rather than physiological outcome measures as the former require no additional hardware or software and are a valid measure of fear conditioning (Boddez et al., 2013). The self-report measures most commonly used in fear conditioning tasks are (1) expectancy ratings (how likely the CS is to be followed by the US during each trial), (2) affective ratings (the level of valence, arousal, or fear felt towards the CS after each phase), and (3) contingency awareness (ability to detect the association between the US and the CS+). These measurements assess different dimensions of fear learning (Lonsdorf et al., 2017). Expectancy ratings and contingency awareness measure explicit knowledge of the associations between stimuli, while affective ratings measure self-reported emotional responses to the stimuli (Baeyens & De Houwer, 1995).

Overview of FLARe

The FLARe software includes a smartphone application and web portal that can be used in tandem to deliver a range of flexible fear conditioning experiments (**Figure 2**). This allows researchers to use one app to deliver numerous different versions of this paradigm. Participants download and use the app to access experiments using a unique login ID, which identifies the participant and the specific experiment in which they are taking part. This login ID serves as the link between app and portal. Researchers use the portal to create and configure their experiments, add or generate participant login IDs, and view/export data from the app. The FLARe portal offers two levels of access to users: admin and researcher. **Admin users** have access to all experiments that exist on the portal whereas

researcher users are only able to view and manage specific experiments and associated data that they created or have been added to.

Experiments are initially created by setting a number of experiment-wide parameters (**Table 1**). For example, researchers are able to upload CS image files and a US audio file to be used throughout the experiment. Once these are set, experiments are built through the use of a number of configurable modules (**Table 2**). Each module is designed with a specific purpose, such as delivering task instructions (task instruction module), or delivering fear conditioning trial screens (fear conditioning module). Each module has modifiable elements according to its purpose. For example, the fear conditioning module allows the number of trials and the rate of reinforcement (the number of times the CS+ is followed by the US) to be determined by the researcher. Modules broadly support three functions; 1) experiment delivery, 2) experiment management, 3) external integration. Modules can be placed anywhere within an experiment and used multiple times if needed. This framework means experiments can range from short one-day experiments that only collect acquisition and extinction data (see **Figure 3**), to longer multi-day experiments that include multiple fear conditioning phases. There are also optional modules to facilitate data collection in-app, such as consenting and screening procedures and delivery of reimbursement via voucher codes after completion, as well as options to add links to additional measures delivered via external survey platforms.

REMOTE FEAR CONDITIONING... WITH FLARe

Create fear conditioning experiments using a web portal and deliver them remotely via an iOS/Android app!

THE SEED

The seed represents **the FLARe codebase** that underlies the FLARe software. Researchers can clone or contribute to the open-source codebase on GitHub.

The open-source codebase is available at <https://github.com/flare-kcl>.

THE POT

The pot represents **the FLARe infrastructure** required to host and deploy the codebase, including web domains, secure servers, and app store accounts.

External parties can collaborate on the existing "FLARe Research" infrastructure or maintain their own infrastructure to host a new instance of the software.

THE PLANT

The plant represents **the FLARe software**. This comprises a web portal researchers can use to create and manage experiments, and a smartphone app that participants engage with to complete the experiment.

This software will be altered if the codebase is changed.

How to use FLARe

LEARN MORE <https://kcl.ac.uk/research/flare>

SOURCE CODE <https://github.com/flare-kcl>

CONTACT US flare@kcl.ac.uk

Figure 2. Infographic for understanding the different aspects of FLARe (codebase, infrastructure, software) and a researcher's journey from creating an experiment to collecting and exporting data.

Table 1. *Global experiment settings.*

Setting	Description
Experiment Code	Maximum six character string associated with the experiment that is used as the prefix for auto-generated participant IDs.
Voucher pool	Researchers can choose to associate their experiment with a pool of voucher codes uploaded to the portal. If an experiment has a voucher pool, a reimbursement screen at the end of the experiment will provide participants with a voucher code.
Expectancy rating anchors	Character strings that are placed at the beginning, mid-point and end of the expectancy rating scales during each trial. These can be changed to reflect your chosen US.
Trial length	The total duration of a fear conditioning trial, defined in seconds.
Rating delay	The delay between the beginning of a fear conditioning trial and the appearance of the expectancy rating scale. This will determine the length of time that the CS and context appear without anything else on the screen.
ITI min/max delay	The lower and upper duration limits for the inter-trial intervals, in seconds. Any one inter-trial interval will last for a random amount of seconds between these upper and lower limits (inclusive).
US	The unconditional stimulus, uploaded as a .wav file.
Minimum device volume	The minimum volume that a participant can lower their phone to without receiving a warning notification. This volume is determined as a proportion of the maximum volume on any given smartphone.
CSA/CSB	The conditional stimuli that will serve as the experiment CS+ (reinforced stimulus) and CS- (unreinforced stimulus). Any .png images that are 800pixel x 800pixel in size can be uploaded. Which image (CSA or CSB) is allocated as the CS+ is randomly determined by the portal for each new participant. Output will be standardised to include both the stimulus name (CSA or CSB) and stimulus type (CS+ or CS-) for each participant.
GS A-D	The four stimuli that will serve as the experiment generalisation stimuli. The GSA should be most similar to the CSA, and the GS D should be most similar to the CSB. When CS+ is allocated for each user, the generalisation stimuli are standardised to match, such that the GS1 will be the most similar to the CS+, and so forth.
Context A-C	Images that will serve as the context image for the experimental trials. Up to three different images can be uploaded as .png files 800pixel x 800pixel. You can later choose to use any one of these context images to serve as the background for fear conditioning trials. Context images are allocated to all trials within a particular fear conditioning phase (e.g. acquisition, extinction, generalisation).

Table 2. Full list of modules available on the FLARe app.

Module	Description	Primary use cases
Affective rating	Allows the researcher to ask a question about CS and GS images (e.g. “how does this image make you feel”). Responses are collected on a nine-point likert scale with modifiable anchors at points one, five, and nine (e.g. “calm”, “neutral”, “anxious”).	Delivering and capturing affective ratings.
Basic information	Captures demographic (date of birth and gender), device (phone make and operating system), and headphone (type: in-ear/on-ear/over-ear) information. Demographic questions can be turned off to maximise anonymity of participants.	Collecting important information about participants and the equipment they are using to take part in the experiment.
Break	Produces two modules, a “break start” and “break end”, for which researchers specify a minimum amount of time that must elapse between them. The two break modules can be placed around other modules to prevent participants from continuing to another section of the experiment too quickly.	Forcing minimum time limits between fear conditioning modules (e.g. between acquisition and extinction phases).
Criterion	Allows researchers to ask any number of yes/no questions (e.g. “do you wish to take part in this study?”). Researchers have full control over question wording and can specify whether the response should be forced (i.e. participants cannot continue without answering). Researchers can also specify a required answer, enabling participants to be screened out if their response excludes them from the study.	Consenting and screening participants before taking part in a study.
Contingency awareness	Asks participants “Did you happen to notice whether you heard the <i>audible keyword</i> after seeing a certain <i>visual keyword</i> ?”. Participants who respond “yes” to the question are taken to a second screen presenting them with an image of both CS to select from. Researchers can modify the audible (e.g. “loud scream”) and visual (e.g. “circle”) keywords so that they match the stimuli used for their specific study.	Determine whether or not participants are contingency aware (i.e. did they recognise which stimulus was the CS+).
Fear conditioning	Delivers a phase of a fear conditioning task (e.g. acquisition or extinction). Researchers can adjust the number of presentations per CS/GS, the number of times the CS+ is reinforced, which context image is used, and whether or not GS are included in the phase.	Delivering different phases of a fear conditioning task, such as acquisition, generalisation, extinction, or renewal.

Setup instructions/calibration	Provides instruction screens that minimally ensure that the smartphone has headphones connected and the device volume is above a limit specified by the researcher. Further instructions can be provided by adding additional screens that allow free text editing. Researchers can also choose to include additional steps to calibrate the volume of the US to a subject point of aversiveness.	Ensuring the participants' device and environment is set up correctly for taking part in the experiment.
Task instructions	Provides instruction screens related to the fear conditioning task itself (e.g. how to make an expectancy rating). The majority of text can be edited and further instructions can be provided by adding additional screens that allow free text editing.	Ensuring the participant knows how to complete the fear conditioning task.
Text	Allows researchers to add additional screens of text that can be formatted using Markdown.	Delivering additional instructions, guidance, or information to participants.
Web	Allows researchers to seamlessly link participants to an external website within the app (e.g. to watch something or undertake questionnaires during a break).	Delivering additional tasks or self-report measures (e.g. through an online survey platform).
Post experiment questions	Asks participants a set of questions about their behaviour during the experiment (e.g. "did you remove your headphones during the task?"). Researchers can choose which, if any, of these questions to ask participants.	Determining task adherence.
US unpleasantness	Asks participants "How unpleasant did you find the <i>audible keyword</i> ?". Researchers can modify the audible (e.g. "loud scream") keyword to match the stimulus used for their specific study. Responses are collected on a ten-point likert scale ranging from "Not unpleasant at all" to "Very unpleasant".	Determining the aversiveness of the US.

Figure 3. Illustration of how modules can be used to build FLARe experiments. All fully flexible modules are listed on the left with an example of how they can be used to create a short experiment displayed on the right. Boxes with a solid border indicate mandatory modules that are always included in the study procedure. Grey boxes indicate modules with a fixed position within an experiment. Arrow to the right indicates a participant's journey through the experiment once they have installed and opened the app.

Architecture

There are three components to FLARe: the codebase, the software, and the infrastructure (**Figure 2**). The FLARe **codebase** contains all the source code required to build the FLARe **software** (the app and portal). The FLARe codebase is open-source, meaning anyone can create their own instance of the software provided they have the **infrastructure** to do so (access to servers, web domain and app stores). The FLARe team based primarily at King's College London have set up their own

infrastructure to create an instance of FLARe known as ‘FLARe Research’. Any future mention of FLARe Research refers exclusively to the FLARe team’s version of the app and portal.

The technical development of the current FLARe software was carried out by a digital project agency, Torchbox (<https://torchbox.com/digital-products/>) to the specifications of, and with active involvement from, the FLARe research team. The FLARe app codebase was written in React Native, a javascript-based open-source mobile application framework that can be used for developing across both Apple and Android platforms to enable maximal participant reach. The FLARe portal was coded using Django, a python-based open-source web framework. This approach facilitates scalable, well documented, easy to read/access code that will enable future rounds of development and the possibility of open-source contributions from external parties. Finally, the portal hosted database is managed using PostgreSQL to facilitate data addition, deletion and searching capabilities, with a readable interface facilitated by the open-source dashboard user-interface kit ‘Tabler’. For detailed technical descriptions of the FLARe suite, see the GitHub hosted code repository (<https://github.com/flare-kcl>), and technical description of the project therein (<https://github.com/flare-kcl/flare-app/blob/develop/paper.md>).

FLARe software is open-source meaning that anybody can set up their own infrastructure for conducting remote fear conditioning research if they wish to. However, in this case, researchers will be required to implement and maintain this infrastructure by themselves. For further technical details please refer to our GitHub code repository <https://github.com/flare-kcl>. Alternatively, researchers may choose to collaborate with the FLARe team at King’s College London to gain access to the ‘FLARe Research’ infrastructure (deployed on a secure Heroku server). The FLARe team is interested in supporting studies that comply with open science practices. As admins of FLARe Research, the FLARe team is able to manage and monitor all experiments being conducted on the platform with collaborating researchers’ access limited to their own experiments.

Remote experimental control

The quality of data collected using a remote fear conditioning task can be adversely affected by participants engaging in a number of unwanted behaviours: (1) removing/unplugging headphones; (2) lowering/varying the intensity of the US by adjusting the phone volume; (3) not engaging with the task (e.g. leaving the app); or (4) completing the task in a non-optimal location or situation. We have included various strategies to discourage these behaviours and track whether or not they occur.

Before the task begins (prevention)

A setup instructions module has been created that ensures headphones are attached and phone volume is appropriate for the study requirements. If these conditions are not met, the app prevents participants from continuing on to the task. The setup instructions module also allows researchers to specify the optimal conditions for completing the task by adding additional instruction screens that allow free text editing. These might, for example, suggest sitting in a quiet room, where the participant will not be disturbed, ensuring their phone is fully charged and access to wifi is available.

During the task (monitoring)

Popup notification flags alert participants if they disconnect their headphones, lower their volume below a researcher's predefined level, or if they fail to make a rating for more than two consecutive trials. During this time, the experiment is paused and cannot continue until the participant acknowledges the notification.

After the task is complete (post-hoc control)

Volume change, failure to make a rating, and minimisation or exiting of the app are logged in participant data for the trial in which they occur. This enables data to be filtered post-hoc based on metrics of participation and engagement. Whilst it is not possible to alert for, or log instances in which participants physically remove their headphones from their ears, this can be assessed via an optional post-experiment questionnaire module. During this, participants can be asked any or all of the following: where the task was completed, if any interruptions were experienced, whether headphones were removed from their ears at any point in the task, and if so, roughly when this occurred.

Applications

FLARe has a number of demonstrable and potential applications, some of which are described below with reference to case studies where applicable.

Making data collection easier

A fundamental application of the FLARe software is to facilitate data collection in large samples, unrestricted by who can physically attend a testing session. As mentioned in the introduction, small sample sizes can lead to underpowered studies with samples that are not representative of the underlying population. The effect of this is amplified in specialist groups (e.g. individuals with

particular diagnoses or exposures), geographically dispersed groups (e.g. nationwide samples), or groups whose access may be restricted (e.g. hospitalised inpatients). With an effective remote experiment, researchers can more easily reach larger, representative samples, and specific groups, than would be possible running an in-person study from a laboratory. FLARe enables this for fear conditioning research, as has been demonstrated via a large-scale ($n > 1,000$) study into anxiety-related differences (McGregor et al., 2021), and preliminary work in specialist groups, including inpatients with an eating disorder diagnosis (Lambert et al., 2021).

Through the collection of large samples, FLARe can also facilitate a change of focus within the field of fear conditioning towards the study of individual differences. A number of criticisms have been levelled at the tendency to study average responding in groups on fear conditioning tasks (Lonsdorf & Merz, 2017). This focus on mean differences could mask underlying patterns of subgroup responding, or poorly describe individual response trajectories (Lonsdorf & Merz, 2017). With an increase in sample sizes, studies will have the power to measure individual differences in responding and gain a more nuanced understanding of different factors associated with fear conditioning. The increase in power will also allow researchers to employ more sophisticated methods of analysis and make use of more complex study designs such as longitudinal, genomic, twin, prediction modelling, and treatment response research. A study using FLARe has already been conducted in a sample of more than 2,000 twins, to examine the heritability of and genetic correlations between fear conditioning outcomes (Purves et al., 2020). Data collection for this study took approximately three months, demonstrating the ease with which large-scale research can be conducted when an existing cohort is employed.

Easy validation of specific study design parameters

Remote studies can be conducted rapidly due to participants being tested simultaneously. The burden on researchers conducting experiments in this way is, therefore, reduced. We hope that if fear conditioning experiments are easier, cheaper and quicker to run, researchers might be encouraged to answer research questions that were previously limited by available resources. This could be particularly useful when attempting to rapidly validate methods or experimental parameters for future work. For example, researchers could easily assess the impact of changes to trial and ITI duration, or the effect of replacing visual (CS) or auditory (US) stimuli. Our team has already taken advantage of the ease with which remote experiments can be conducted to validate a previous version of the FLARe app against an in-person version of the task (Purves et al., 2019).

Discussion

Summary

As stated in the rationale, we aimed to create a platform for assessing fear conditioning remotely that would improve data sharing and methodological standardisation. Both these aims have been met in the latest round of development, with scope to refine and develop the research tool further.

The FLARe app and web portal provide a remote alternative to in-laboratory fear conditioning, enabling a genuine advancement in the scale of fear conditioning data collection. To date, the FLARe app has been used to collect data from a total of over 2,750 individuals. Data from approximately 2,500 of these participants was collected in just three months, supported for the majority of that time by just one team member. In comparison, our prior pilot and validation work, using in-person testing, collected data from 231 participants in 15 months, using at least three full-time team members. For context, our total sample collected via the FLARe app is greater than the largest meta-analysis of fear conditioning (Duits et al., 2015), which includes 2,185 participants. The largest individual study within this analysis included 270 participants (Norrholm et al., 2011).

FLARe also represents an opportunity to improve the accessibility - to both participants and researchers - and standardisation of fear conditioning across studies. Accessibility has been achieved by creating a Graphical User Interface (GUI) web portal that makes it easy for researchers to set up and track their fear conditioning experiment without any special coding or software skills. The web portal facilitates standardisation by only allowing the customisation of parameters that commonly vary across fear conditioning studies (e.g. reinforcement rate, trial duration, CS image or US sound), whilst holding other aspects of the task constant (e.g. visual design and user experience).

Additionally, all parameter choices are tracked via the portal meaning that, in the future, if multiple research teams use the same infrastructure, it should be simple to identify studies using similar designs. Similarity across studies will be increased due to the inherent consistencies within the app, mostly owing to standardised user experience. This will broaden the potential for multi-group collaboration, as well as meta- and mega-analyses. The numerous applications of the FLARe app, such as collecting large, representative samples, facilitating novel study designs, and making it easier to validate extensions of the paradigm, should all help to expand on and improve the reproducibility of the field.

Limitations

The fundamental disadvantage of conducting remote research is the inability to control all aspects of the testing environment and participant behaviour that would normally be limited or observed by the experimenter in the laboratory. For example, location, time of day, and external distractions may all interfere with the participant's behaviour during the experiment. A number of strategies can be employed to reduce the likelihood and detect the occurrence of such behaviours using FLARe (see "Remote experimental control" section above). Additionally, studies have shown that results between in-laboratory and remote versions of experiments are comparable (Casler, Bickel, & Hackett, 2013; Purves et al., 2019), which suggests that these limitations are unlikely to have a significant impact on research outcomes when data is obtained remotely.

Another limitation of remote data collection is the greater risk of drop-out after starting the task. Because it is easier to reach more people, more rapidly, it is likely that the number of participants who complete the app will still be greater than the number that could have been obtained with complete data in a comparable laboratory set-up. More concerning, however, is that certain groups of participants may drop out of the experiment due to important characteristics, such as greater levels of anxiety (McGregor et al., 2021). These participants may be more likely to continue the task with a researcher present. To reduce the likelihood of dropout due to the aversiveness of the fear conditioning task, the FLARe team have developed a number of features for researchers using the FLARe app. These features include the ability to use an alternative aversive stimulus as the US and a module for participants to calibrate the volume at which the US is played. There is some evidence to suggest that the default US used to validate the FLARe app (a loud female scream; Purves et al., 2019) can result in a large proportion of dropout compared to other stimuli that are rated as equally aversive (Shechner et al. 2015; Michalska et al. 2016). Use of an alternative US may reduce the number who attempt to avoid the sound by removing their headphones or lowering the volume of their phone. Furthermore, sensitivity to the US varies amongst participants. Whilst the validated version of the FLARe app (Purves et al., 2019) forced all participants to set their phone volumes to 100%, the latest version includes an optional experimental module that allows each participant to calibrate the US based on their own subjective ratings of aversiveness. In this module, participants hear the US at different volumes, selecting a volume that they personally find to be aversive, but not intolerable. Using a calibration module to standardise the volume based on subjective ratings of aversiveness rather than absolute volume may also improve the tolerance of the US and reduce dropout. Whilst this approach reflects the calibration usually used for the electric shock that is often

the US for in-person fear conditioning, it is important to note that this new feature is yet to be validated against in-person testing.

Future development

There are several development plans for progressing and improving the FLARe software. These proposals will be developed to work alongside the FLARe Research infrastructure, but will depend on funding and the priority of features for planned studies.

Templates and reproducibility

At present, researchers from different groups need to communicate closely to ensure that their paradigms are comparable. In future rounds of development, we hope to improve standardisation across studies by introducing user generated templates for sharing experiment design, setup and materials (e.g. image and sound files) with other groups in a way that is immediately replicable.

Centralised and open database

FLARe's ability to collect big data on fear conditioning could be further improved by refining the web portal's database. As it stands, certain portal users (admins) are able to view data from all studies conducted using a specific instance of the FLARe software, an arrangement that needs to be considered in ethical approval and data sharing agreements with collaborating groups. For FLARe Research, this is restricted to the FLARe team admins who require access to all studies to facilitate successful use of the infrastructure by collaborating groups. It would be theoretically possible, given appropriate informed consent and ethical agreement, for admins to combine and analyse comparable anonymised data collected across different studies and geographic regions. Researchers collaborating on the same infrastructure (e.g. FLARe Research) would, therefore, be able to combine datasets to improve sample size further. We hope to develop this capability further to create a formalised central database where fully anonymised versions of data from studies collected on the same infrastructure could be combined, viewed and exported together. This data would then be accessible to any researchers upon submission of an appropriate secondary data request that demonstrated a clear hypothesis, research plan, and ethical approval. Preparatory work for this future development includes careful consideration of variables collected using the FLARe app, ensuring that little to no identifiable data is stored on the database, and that all studies conducted using the FLARe Research infrastructure meet appropriate ethical standards and are compliant with data processing requirements. Future research conducted using the FLARe Research infrastructure

will be required to comply with open science practices so that our centralised database of fully anonymised data can be shared openly, with due consideration of data protection and the rights of participants.

Incorporating measurements across multiple modalities

As mentioned previously, emotional responses to fear conditioning can be measured on a number of levels (e.g. self-report, behavioural, physiological, or neurobiological; see Lonsdorf et al. 2017 for a review). FLARe's reliance on self-report measures of fear conditioning is, therefore, one limitation we hope to address in the future. Aside from only measuring one level of 'fear', self-report outcome measures have been criticised for their susceptibility to subjective biases and their lack of direct comparison with animal research (Lonsdorf et al. 2017). Incorporating additional measures that tap into levels of fear that are not susceptible to subjective biases (e.g. physiological measures such as skin conductance response, fear potentiated startle, or heart rate) is a priority for future rounds of development. The choice to adopt smartphone technology gives FLARe the advantage of being able to incorporate physiological measures by making use of the phone's built-in sensors, while continuing to conduct research remotely. At this point, a heart rate measure is the most likely candidate for integration into the app, since it can be collected in a number of different ways (e.g. through the smartphone's camera/light or via wearables; Bai, Hibbing, Mantis, & Welk, 2018; Coppetti et al., 2017; Hwang et al., 2019). We hope to continue expanding the number of measures available to the researcher, allowing the measurement of different levels and components of fear.

Collaboration statement

Recent improvements to FLARe and its underlying codebase now mean that it is ready to be used by other researchers through open-source sharing of the software or collaboration with the FLARe team using the FLARe Research infrastructure. The FLARe team aims to improve reproducibility within the field of fear conditioning research by forming collaborations that will lead to large-scale multi-site studies of fear conditioning in line with a manifesto for reproducible science (Munafò et al., 2017). The FLARe team is open to discussing models for collaboration with interested groups, and encourages all enquiries or questions.

Concluding remarks

The FLARe smartphone app and web portal provides researchers the unique opportunity to conduct fear conditioning research remotely. These tools have been created with other researchers in mind, prioritising flexibility in study design while limiting other aspects of customisation for ease of use. Without the usual barriers associated with laboratory testing, researchers can use FLARe to carry out large scale, comparable fear conditioning studies with easily reproducible methodologies. FLARe holds the potential to collect 'big data' that could open the field of fear conditioning research to unprecedented sample sizes, offering enhanced reproducibility, reliability and the power to examine nuanced individual differences in individuals with anxiety disorders.

Declarations:

Funding

T. McGregor is supported by the UK Medical Research Council (MR/N013700/1). T.C. Eley and G. Breen are part-funded by a program grant from the UK Medical Research Council (MR/M021475/1 and MR/V012878/1). This study presents independent research part-funded by the Wellcome Trust (221582/Z/20/Z) and National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Code availability

<https://github.com/flare-kcl>

Authors' contributions

KLP/TCE/GB originally conceived remote fear conditioning by app. TM/KLP/TCE conceived a flexible and multipurpose tool for conducting fear conditioning experiments remotely and collaboratively. TB/EC/MGC/KSY/TCE/GB provided technical advice and oversight throughout the development process. TM/KLP co-ordinated app development and wrote the manuscript. TCE/KSY gave advice and feedback at several stages of manuscript drafting. All authors reviewed the manuscript.

Acknowledgements

The authors would like to acknowledge the many people and groups who have made this work possible, or who's contributions have improved it over the course of development. Torchbox Ltd, in particular Ben Heasman, Nathan Horrigan, Mitchel Cabuloy, Paul Vetch. The Twins Early Development research team and participants. Members of the Social, Genetic and Developmental Psychiatry Centre of King's College London who have contributed to validation and testing of the app; Omowonuola Akingbuwa, Zara Ampong-Appiah, Megan Hammond Bennett, Michela Clarke, Molly Davies, Sarah Derveeuw, Bernadette De Villiers, Laurie Hannigan, Jenny Leng, Bridie MacDonald, Ellie Minshall, Dina Monssen, Elisavet Palaiologou, Shivani Parikh, Alicia Peel, Jade Pusey, Katie White and the EDIT Lab team. Contributors to early feasibility testing in UCLA; Michael Sun, Michael Treanor. The many collaborators and scientists with whom we have had informal conversations by email or at conferences over the years, many of which have helped shape our thinking.

Conflicts of interest/Competing interests

None to declare

Ethics approval

Not applicable

Consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Not applicable

References

- Adjerid, I., & Kelley, K. (2018). Big data in psychology: A framework for research advancement. *The American Psychologist*, 73(7), 899–917. <https://doi.org/10.1037/amp0000190>
- Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, 52(1), 388–407. <https://doi.org/10.3758/s13428-019-01237-x>
- Bai, Y., Hibbing, P., Mantis, C., & Welk, G. J. (2018). Comparative evaluation of heart rate-based monitors: Apple Watch vs Fitbit Charge HR. *Journal of Sports Sciences*, 36(15), 1734–1741. <https://doi.org/10.1080/02640414.2017.1412235>
- Boddez, Y., Baeyens, F., Luyten, L., Vansteenwegen, D., Hermans, D., & Beckers, T. (2013). Rating data are underrated: validity of US expectancy in human fear conditioning. *Journal of Behavior Therapy and Experimental Psychiatry*, 44(2), 201–206. <https://doi.org/10.1016/j.jbtep.2012.08.003>
- Brown, H. R., Zeidman, P., Smittenaar, P., Adams, R. A., McNab, F., Rutledge, R. B., & Dolan, R. J. (2014). Crowdsourcing for cognitive science--the utility of smartphones. *Plos One*, 9(7), e100662. <https://doi.org/10.1371/journal.pone.0100662>
- Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 29(6), 2156–2160. <https://doi.org/10.1016/j.chb.2013.05.009>
- Coppetti, T., Brauchlin, A., Müggler, S., Attinger-Toller, A., Templin, C., Schönrath, F., ... Wyss, C. A. (2017). Accuracy of smartphone apps for heart rate measurement. *European Journal of Preventive Cardiology*, 24(12), 1287–1293. <https://doi.org/10.1177/2047487317702044>
- Cortis Mack, C., Harding, M., Davies, N., & Ward, G. (2019). RECAPP-XPR: A smartphone application for presenting and recalling experimentally controlled stimuli over longer timescales. *Behavior Research Methods*, 51(4), 1804–1823. <https://doi.org/10.3758/s13428-018-1157-x>
- Craske, M. G., Hermans, D., & Vervliet, B. (2018). State-of-the-art and future directions for extinction

- as a translational model for fear and anxiety. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 373(1742). <https://doi.org/10.1098/rstb.2017.0025>
- de Leeuw, J. R. (2015). jsPsych: a JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, 47(1), 1–12. <https://doi.org/10.3758/s13428-014-0458-y>
- Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., ... Baas, J. M. P. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety*, 32(4), 239–253. <https://doi.org/10.1002/da.22353>
- Dymond, S., Dunsmoor, J. E., Vervliet, B., Roche, B., & Hermans, D. (2015). Fear generalization in humans: systematic review and implications for anxiety disorder research. *Behavior Therapy*, 46(5), 561–582. <https://doi.org/10.1016/j.beth.2014.10.001>
- Gureckis, T. M., Martin, J., McDonnell, J., Rich, A. S., Markant, D., Coenen, A., ... Chan, P. (2016). psiTurk: An open-source framework for conducting replicable behavioral experiments online. *Behavior Research Methods*, 48(3), 829–842. <https://doi.org/10.3758/s13428-015-0642-8>
- Hwang, J., Kim, J., Choi, K. J., Cho, M. S., Nam, G. B., & Kim, Y. H. (2019). Assessing Accuracy of Wrist-Worn Wearable Devices in Measurement of Paroxysmal Supraventricular Tachycardia Heart Rate. *Korean Circulation Journal*, 49(5), 437–445. <https://doi.org/10.4070/kcj.2018.0323>
- Jardine, J., Fisher, J., & Carrick, B. (2015). Apple's ResearchKit: smart data collection for the smartphone era? *Journal of the Royal Society of Medicine*, 108(8), 294–296. <https://doi.org/10.1177/0141076815600673>
- Lambert, E. R., Treasure, J., Purves, K. L., McGregor, T., Bergou, N., Kan, C., ... Cardi, V. (2021). Fear Conditioning in Women With Anorexia Nervosa and Healthy Controls: A Preliminary Study. *Journal of Abnormal Psychology*.
- Lange, K., Kühn, S., & Filevich, E. (2015). “just another tool for online studies” (JATOS): an easy solution for setup and management of web servers supporting online studies. *Plos One*, 10(6), e0130834. <https://doi.org/10.1371/journal.pone.0130834>
- Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S.

- (2005). Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behaviour Research and Therapy*, 43(11), 1391–1424. <https://doi.org/10.1016/j.brat.2004.10.007>
- Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., ... Merz, C. J. (2017). Don't fear "fear conditioning": Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience and Biobehavioral Reviews*, 77, 247–285. <https://doi.org/10.1016/j.neubiorev.2017.02.026>
- McGregor, T., Purves, K. L., Constantinou, E., Baas, J. M. P., Barry, T. J., Carr, E., ... Eley, T. C. (2021). Large-scale remote fear conditioning: Demonstration of associations with anxiety using the FLARe smartphone app. *Depression and Anxiety*. <https://doi.org/10.1002/da.23146>
- Mineka, S., & Oehlberg, K. (2008). The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. *Acta Psychologica*, 127(3), 567–580. <https://doi.org/10.1016/j.actpsy.2007.11.007>
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., du Sert, N. P., ... Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1, 0021. <https://doi.org/10.1038/s41562-016-0021>
- Norrholm, S. D., Jovanovic, T., Olin, I. W., Sands, L. A., Karapanou, I., Bradley, B., & Ressler, K. J. (2011). Fear extinction in traumatized civilians with posttraumatic stress disorder: relation to symptom severity. *Biological Psychiatry*, 69(6), 556–563. <https://doi.org/10.1016/j.biopsych.2010.09.013>
- Pew Research Center. (2021). *Demographics of Mobile Device Ownership and Adoption in the United States*. Pew Research Center.
- Purves, K. L., Constantinou, E., McGregor, T., Lester, K. J., Barry, T. J., Treanor, M., ... Eley, T. C. (2019). Validating the use of a smartphone app for remote administration of a fear conditioning paradigm. *Behaviour Research and Therapy*, 123, 103475. <https://doi.org/10.1016/j.brat.2019.103475>
- Purves, K. L., Krebs, G., McGregor, T., Constantinou, E., Lester, K. J., Barry, T. J., ... Eley, T. C. (2020). Evidence for distinct genetic and environmental influences on fear acquisition and extinction.

BioRxiv. <https://doi.org/10.1101/2020.12.18.423504>

Sochat, V. V., Eisenberg, I. W., Enkavi, A. Z., Li, J., Bissett, P. G., & Poldrack, R. A. (2016). The experiment factory: standardizing behavioral experiments. *Frontiers in Psychology*, 7, 610.

<https://doi.org/10.3389/fpsyg.2016.00610>

Vagell, R., Vagell, V. J., Jacobs, R. L., Gordon, J., & Baden, A. L. (2019). SMARTA: Automated testing apparatus for visual discrimination tasks. *Behavior Research Methods*, 51(6), 2597–2608.

<https://doi.org/10.3758/s13428-018-1113-9>