THE UNIVERSITY of York

Application Form for Physical Sciences Ethics Committee Approval

Advice for applicants on completing the form

Please ensure that the information provided is:

- *Accurate and concise*
- *Clear and simple and easily understood by a lay person*
- *Free of jargon, technical terms and abbreviations*

Further advice and information can be obtained from your departmental representative on the PSEC and at: http://www.york.ac.uk/admin/aso/ethics/cttee.htm

Please return completed form to your departmental representative:

Daniel Kudenko, Department of Computer Science

Title of project: Enjoyment and data quality in a human-subject data collection game

SECTION 1 DETAILS OF APPLICANTS

Details of principal investigator (name, appointment and qualifications)

David Gundry, PhD Student, Department of Computer Science, University of York

Names, appointments and qualifications of additional investigators *(student applicants should include their project supervisor(s) here)*

Sebastian Deterding, Reader in the Department of Theatre, Film and Television, University of York.

Location(s) of project

Online

SECTION 2 FUNDERS

What is the funding source(s) for the project?

The EPSRC Centre for Doctoral Training in Intelligent Games & Games Intelligence (IGGI) [EP/L015846/1] and the Digital Creativity Labs (digitalcreativity.ac.uk), jointly funded by EPSRC/ AHRC/Innovate UK under grant no. EP/M023265/1.

Please answer the following:

(i)	Does the express and direct aim of the research or other activity raise ethical issues?	YES		NO	X
(ii)	Is there any obvious or inevitable adaptation of research findings to ethically questionable aims?	YES		NO	X
(iii)	Is the work being funded by organisations tainted by ethically questionable activities?	YES		NO	X
(iv)	Are there any restrictions on academic freedoms – notably, to adapt and withdraw from ongoing research, and to publish findings	YES		NO	X

If you answered *Yes* to any of the above, please give details below:

SECTION 3 DETAILS OF PROJECT OR OTHER ACTIVITY

Aims (100 words max)

This research is about meta-methodological issues regarding data collection with applied games. In this context an applied game is a game that is designed specifically for collecting a certain type of scientific data.

An applied game has been developed to evaluate how game design affects the validity of data that is collected. In this experiment we compare a game and non-game methodology for collecting a specific type of data: adjective word-order. We investigate effects on enjoyment and data quality.

A theoretical model and set of design principles for the design of games for subjective data collection has been proposed (http://eprints.whiterose.ac.uk/133118/). This work is the first in a series of experiments which are planned using this game to evaluate this model.

Background (250 words max)

Games are increasingly being used as data collection tools in academic research. They have the potential to make engaging with experimental research more enjoyable for participants. However, not enough is known about how the use of games might themselves affect the quality of the data that is collected. This is a significant problem in domains where the data cannot be externally validated, particularly human-subject data such as linguistic performance.

A human-subject data collection game is an applied game that collects data about the player, such as their patterns of language use. Games have been used in research studies about e.g. error rates in number entry. However, a class of data which I call subjective data (e.g. what is your favourite ice-cream flavour) has not been collected with games. This class of data is useful for e.g. linguistic science, which uses subjective data about languistic practice.

Adjective word order is the relative positions of different types of adjectives in a phrase ("big red square" is grammatical in English where "red big square" is not). This varies between languages but is largely consistent between speakers of the same language. It is an instance of linguistic human-subject data which can in principle be subjective to the individual but allows us to use the 'standard' form as a ground truth for comparison.

Brief outline of project/activity (250 words max)

Participants will play a digital game. This game will be accessed online via a web browser. Participants will arrive at the game via Prolific (https://www.prolific.co), a research participant panel provider. The study description and consent information will be presented in the experiment description provided to participants on Prolific, along with details of payment. By choosing to begin the experiment, participants indicate their consent.

Participants will be directed to a game or non-game task implemented using the same interface. This game will present brief instructions and some demographic questions to the participant. Participants will then play through a tutorial. Afterwards they will then play the game until they have made 20 inputs (exlcuding inputs made in the tutorial). Then the game will end and participants will be asked to complete a short questionnaire about their experience. They w ill then be thanked for their time and directed back to the Prolific platform.

The game requires players to identify shapes by clicking on words (adjectives and nouns). Selecting a 'sentence' of three words counts as one input (this sentence must contain exactly one noun). Selecting shapes in this way causes them to be cleared from the screen. The remaining blocks fall down to fill the space. Points and bonus moves are gained depending on the number and arrangement of blocks cleared. The goal is to clear the screen without running out of moves.

In the other condition, players are asked to describe an indicated shape on the screen using the same input method. Points, moves and other game features are removed.

Study design *(if relevant – e.g. randomised control trial; laboratory-based)*

Online double-blind between groups design

If the study involves participants, how many will be recruited? 100

If applicable, what is the statistical power of the study, i.e. what is the justification for the number of participants needed? See the attached file `power-analysis.Rout`.

SECTION 4 RECRUITMENT OF PARTICIPANTS

How will the participants be recruited?

- Participants will be recruited from Prolific (https://www.prolific.co/), a panel provider for research studies.

What are the inclusion/exclusion criteria?

- Participants will be pre-filtered by Prolific to include only adults (18+) whose first language is English.
- Participants who report their age as under 18 will be excluded and payment will be withheld.

YES		NO	X

YES	X	NO	

Will participants be paid reimbursement of expensesWill participants be paid?If yes, please obtain signed agreementWill any of the participants be students?

Due to the recruitment process, I have no way of knowing.

SECTION 5 DATA STORAGE AND TRANSMISSION

If the research will involve storing personal data, including sensitive data, on any of the following please indicate so and provide further details (answers only required if *personal* data is to be stored).

Manual files	
University computers	
Home or other personal computers	
Laptop computers, tablets	
Website	Restdb.io database

Please explain the measures in place to ensure data confidentiality, including whether encryption or other methods of anonymisation will be used.

The data collected can only be used to identify an individual through their 'Prolific ID'. And only then in combination with Prolific's confidential database. The Prolific ID is required for tracking participant completion and payments. For all other uses, the data can be anonymised by removing the Prolific ID.

Data will be submitted by the user's browser to a Restdb.io (https://restdb.io) database via (encrypted) HTTPS. David is the only person with access to this database. It will be stored in plaintext until it is downloaded (again over HTTPS) to David's computer. It will then be deleted from the database.

It will be stored are processed on David's university computer. It will be anonymised and subsequently stored and processed on David's personal computers (home desktop and laptop), as well as those of his supervisor. This anonymised data will be published for research transparency on the Open Science Framework (https://osf.io)

Please detail who will have access to the data generated by the study.

David Gundry. Anonymised data will be published on the Open Science Framework (https://osf.io) and shared with others. Data will not be aggregated before publication, each participant's data will be associated with a unique number.

Please detail who will have control of and act as custodian for, data generated by the study.

David Gundry

Please explain where, and by whom, data will be analysed.

Data will be anonymised before analysis. It will be analysed by David Gundry.

Please give details of data storage arrangements, including where data will be stored, how long for, and in what form.

Data will be stored on restio.db until data collection is complete. It will be stored on David's university computer in non-anonymised form until no more than a month after all participants have been paid. Afterwards it will be stored only in anonymised form.

SECTION 6 CONSENT

Is written consent to be obtained?

YES		NO	X

If yes, please attach a copy of the information for participants

If no, please justify

Detailed study and consent information will be included in the study description on Prolific. By clicking to begin the study, they indicate their consent. The description provided to participants is supplied in the attached file `description.md`.

Will any of the participants be from one of the following vulnerable groups?

Children under 18	YES		NO	X
People with learning difficulties	YES		NO	X
People who are unconscious or severely ill	YES		NO	X
People with mental illness	YES		NO	X
NHS patients	YES		NO	X
Other vulnerable groups (if 'yes', please give details)	YES		NO	X

If so, what special arrangements have been made for getting consent?

While no vulnerable groups are specifically targeted (and anyone under 18 is excluded by Prolific), I have no way of knowing whether a member of a vulnerable group participates.

SECTION 7 DETAILS OF INTERVENTIONS

Indicate whether the study involves procedures which:

Involve taking bodily samples	YES		NO	X
Are physically invasive	YES		NO	X
Are designed to be challenging/disturbing (physically or psychologically)	YES		NO	X

If so, please list those procedures to which participants will be exposed:

List any potential hazards: None beyond everyday casual gameplay.

List any discomfort or distress: None

What steps will be taken to safeguard

(i) the confidentiality of information

Information is only personally identifiable in combination with Prolific's confidential data. This will be securely stored only as long as needed and then destroyed.

(ii) the specimens themselves?

What particular ethical problems or considerations are raised by the proposed study?

None.

What do you anticipate will be the output from the study? ***Tick those that apply:***

Peer-reviewed publications	X
Non-peer-reviewed publications	X
Reports for sponsor*	
Confidential reports	
Presentation at meetings	X
Press releases	

* If you have ticked this box, then please identify the sponsor ……

……

Is there a secrecy clause to the research?	YES		NO	**X**

If yes, please give details below

SECTION 8 SIGNATURES

The information in this form is accurate to best of my knowledge and belief and I take full responsibility for it.

I agree to advise of any adverse or unexpected events that may occur during this project, to seek approval for any significant protocol amendments and to provide interim and final reports. I also agree to advise the Ethics Committee if the study is withdrawn or not completed.

Signature of Investigator(s): David Gundry……

……

Date: ……

Responsibilities of the Principal Researcher following approval

- If changes to procedures are proposed, please notify the Ethics Committee
- Report promptly any adverse events involving risk to participants