ROYAL INSTITUTE
OF TECHNOLOGY

DD1365
SOFTWARE ENGINEERING

USER STUDIES

Christopher Peters

HPCViz, KTH Royal Institute of Technology, Sweden

chpeters@kth.se

http://kth.academia.edu/ChristopherEdwardPeters

Who am I?

Christopher Peters

Associate Prof.

Research:

Virtual characters, computer graphics, computer games, perceptual computing

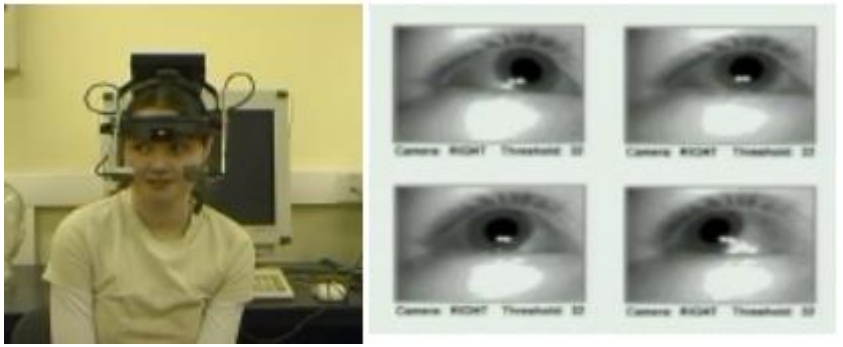
Teaching:

DD3336 Interactive Entertainment Technologies (PhD level)

DH2650 Computer Game Design

DH2323 Computer Graphics

C. Peters and C. Ennis. Modeling groups of plausible virtual pedestrians, Computer Graphics and Applications, IEEE 29 (4), 54-63, 2009



C. Peters and C. O'Sullivan. Attention-driven eye gaze and blinking for virtual humans, ACM SIGGRAPH 2003 Sketches & Applications, 2003

ROYAL INSTITUTE OF TECHNOLOGY

Annoying Humans

- Modern software engineering inherently **human-centered**
- Requirements, architecture, design, development, testing, maintenance, processes, outputs
- **Primarily created, evaluated and performed by humans**

ROYAL INSTITUTE OF TECHNOLOGY

Annoying Humans (II)

- Major stakeholders
  - Developers
  - Managers
  - Consumers
- **Primarily humans too!**

ROYAL INSTITUTE OF TECHNOLOGY

Obvious?

- Blindingly!

- But it takes exceptional and continuous conscious effort to *properly* keep humans in the process
- Many pitfalls
- Need an experimental methodology to help

ROYAL INSTITUTE OF TECHNOLOGY

User Studies for Evaluation

- Human experiments
- Process of evaluating or understanding a technique, tool or idea in terms of needs, preferences and abilities of humans
- Have people use your system or observe stimuli
- Evaluate what they do
- Need for computer scientists to conduct user studies

ROYAL INSTITUTE OF TECHNOLOGY

“Can’t *I* just do it?”

- As a developer of a system, you…
  - Are a system expert
  - Have insider knowledge
  - Are familiarized with the system
  - Do not have the perspective of a user

- **You are inherently biased**

  You are a very poor adviser for the design, development or evaluation of your system

ROYAL INSTITUTE OF TECHNOLOGY

User Studies

- Covers a broad range of activities involving users
- User design
- Usability
- Psychophysics experiments

ROYAL INSTITUTE OF TECHNOLOGY

Process Overview

- Design
- Procedure
- Data analysis
- Conclusions

ROYAL INSTITUTE OF TECHNOLOGY

Process Overview

- Design
  - *Hypothesis*: what do you want to find out?
  - Who will be the *population*?
  - *Metrics*: what will be measured / recorded?

ROYAL INSTITUTE OF TECHNOLOGY

Process Overview

- Design
- Procedure
  - All participants sign up for a time slot
  - *Informed consent*
  - Execute study
  - Questionnaires/debrief

ROYAL INSTITUTE OF TECHNOLOGY

Process Overview

- Design
- Procedure
- Data analysis
  - Chance and confidence: *Significance*
  - *T-test*
  - *ANOVA*
  - F statistic, *p* values

ROYAL INSTITUTE OF TECHNOLOGY

The Role of Chance

85% success rate: Euro 2008, World Cup 2010

Paul the Octopus, Animal Oracle

ROYAL INSTITUTE OF TECHNOLOGY

Response Measures

- Free description
- Rating scales
- Forced-choice
- Multiple choice
- Real-world tasks

- Choose wisely…

ROYAL INSTITUTE OF TECHNOLOGY

Task-based Studies

- **Comparative**

  Perform same task under different conditions
- **Observation**

  Observed performing a specific task
- **Field**

  Observations of people in the field performing various activities
- Time-consuming: one must design and observe tasks

ROYAL INSTITUTE OF TECHNOLOGY

Controlled Experiments

- Events or actions caused by the experimenter intentionally
- Controlled: only variables being examined will change

  Everything held constant except for one variable
- Control group: *normal* or *usual* state
- Repeatedly and reliably produce a specific event or situation

Cause and effect (correlation v causation)

ROYAL INSTITUTE OF TECHNOLOGY

The Task

- Clearly specify it (requires effort)
- Evaluation:

  “A mouse is faster than a keyboard for numeric entry”
- Hypothesis:

  “Participants using a keyboard to enter a string of numbers will take less time than participants using a mouse”

ROYAL INSTITUTE OF TECHNOLOGY

Variables

- *Independent variable*: manipulated by the experimenter
- *Dependent variable*: caused by the independent variable
- In the previous example
  - Independent variable: interaction method (mouse / keyboard)
  - Dependent variable: time

ROYAL INSTITUTE OF TECHNOLOGY

Conditions

- Each condition changes something
- Independent variables (IV)

- In controlled experiment:
  - Two group types: *Control group* and *Experiment group(s)*

- Need to consider the ordering of conditions

ROYAL INSTITUTE OF TECHNOLOGY

Participants

- *Within-subjects* vs *between-subjects*
- Within-subjects
  - *Repeated measures* design
  - Participant tested under each condition

ROYAL INSTITUTE OF TECHNOLOGY

Participants

- *Within-subjects* vs *between-subjects*
- Within-subjects
- Between-subjects
  - *Independent measures*
  - Participant tested under one condition only
  - Avoid order effects, boredom; more participants needed

ROYAL INSTITUTE OF TECHNOLOGY

Participants

- Record *relevant* participant details!
  - Gender
  - Age
  - Handedness
  - Vision
- Pay close attention to ethics/legal considerations!
  - **Anonymity**
    - Data needs *to be* anonymous and participant needs to *know*

ROYAL INSTITUTE OF TECHNOLOGY

Notes

- Power: the more participants there are, the better they sample the population
- ~20 participants per condition often considered a minimum number
- But finding volunteers for user studies is difficult…

  Even if you pay them

ROYAL INSTITUTE OF TECHNOLOGY

Free description

- Participants asked for beliefs and opinions
- Questions in text form
- Qualitative

- **Interviews**
- **Questionnaires**
- **Long/short answers**

ROYAL INSTITUTE OF TECHNOLOGY

Think aloud

- Participants ‘think aloud’ as they perform a set of specified tasks
- Describe what they are looking at, thinking, doing and feeling during the task
- Observers take notes
- Observers see first hand the process of task completion
  - Not just final product

ROYAL INSTITUTE OF TECHNOLOGY

The Test Environment

ROYAL INSTITUTE OF TECHNOLOGY

The Test Environment

ROYAL INSTITUTE OF TECHNOLOGY

Pitfall #1

People sometimes do strange things, so they need to be observed

ROYAL INSTITUTE OF TECHNOLOGY

Pitfall #2

People sometimes do strange things because they are being observed

ROYAL INSTITUTE OF TECHNOLOGY

Pitfall #3

- Be very careful about the wording of questions

  "About how fast were the cars going when they *smashed* into each other?"

  (Loftus & Palmer, 1974)

- **Garbage in -> garbage out**

ROYAL INSTITUTE OF TECHNOLOGY

Pitfall #4

- Experimenter bias
- **Seeks evidence conforming to one's expectations**
- 'Cherry picking'
  - Keep/focus on the *good* data, discard/ignore *bad* data
- Unintentional
- There are *many* more

  Google: "List of cognitive biases"

ROYAL INSTITUTE OF TECHNOLOGY

Pitfall #5

- Response bias
- **Participants may try to give you the answers they think you want**
- Conceal expectations
- Preserve anonymity
  - Data collection should be anonymous
- Add *catch trials*

ROYAL INSTITUTE OF TECHNOLOGY

General Advice

- Always do a *pilot study*
- Smaller number of participants
- Not statistically valid
- But highlights problems with the experiment design and procedure…

  …***before*** the main experiment

ROYAL INSTITUTE OF TECHNOLOGY

A ‘Live’ Example

ROYAL INSTITUTE OF TECHNOLOGY

4 Experiment

Thirty two participants (12F, 20M) age 18 to 30, were seated in front of a computer screen. They were told that the experiment consists of three blocks and were given an instruction sheet: two photographs of the corridor and open zone were shown and they were told that the images they were about to see were derived from real photographs, but in some the character formations were real, while in others they were synthetically generated. For the first block of the experiment the participants were told to focus only on the positions of the characters. For each image displayed, participants were asked if they thought the positions of the *pawn* figure characters were real or synthetically generated. For the second block, participants were asked to look at the orientations of the characters only and judge if they were real or synthetically generated. For the final block of the experiment, participants were asked to take both position and orientation of the characters into account and judge whether the scenes were real or synthetically generated. The reason that we presented the experiment in this order was to avoid biasing participants. If the pawn figures were viewed after the humanoid characters, this could have caused them to perceive the scenes as less realistic due to the reduced realism of the characters, which was not the effect being tested. Furthermore, the scenes with position and orientation combined were presented during the final block, to prevent participants from taking position into consideration when conducting the orientation only trial. Between each trial, a blank-screen was displayed for 5 seconds, after which the number of the next trial was displayed alerting participants.

Cathy Ennis, Christopher Peters, Carol O'Sullivan: Perceptual evaluation of position and orientation context rules for pedestrian formations. Applied Perception in Graphics and Visualization (APGV) 2008: 75-82

Methodology

- Consisted of 4 phases:
  - Data Collection Phase
  - Annotation Phase
  - Reconstruction Phase
  - **Modification Phase**

Data Collection Phase

- Videos taken of 2 locations:

Unconstrained / Open Scene
30 Characters

Constrained / Corridor Scene
12 Characters

APGV 08
applied perception in graphics and visualization

Annotation Phase

- Still images annotated to highlight Positions, Orientations and Groupings
- Colour-coded *Dynamic vs. Static* groups and *8 different Orientations*

APGV 08
applied perception in graphics and visualization

ROYAL INSTITUTE OF TECHNOLOGY

Introduction **Methodology** Experiment Results Conclusions

Position Rules

Still Image

Real

Random

Context:
Bounds Sensitive, Group Sensitive

APGV 08
applied perception in graphics and visualization

Orientation Rules

Still Image

Real

Random

Context:
Flow Sensitive, Adjacency Sensitive, Group Sensitive

APGV 08
applied perception in graphics and visualization

Reconstruction Phase

- Creation of virtual replicas of real images that were captured and annotated

- Using image as viewport background in 3ds Max
- Tweaking Camera parameters to align model and still image

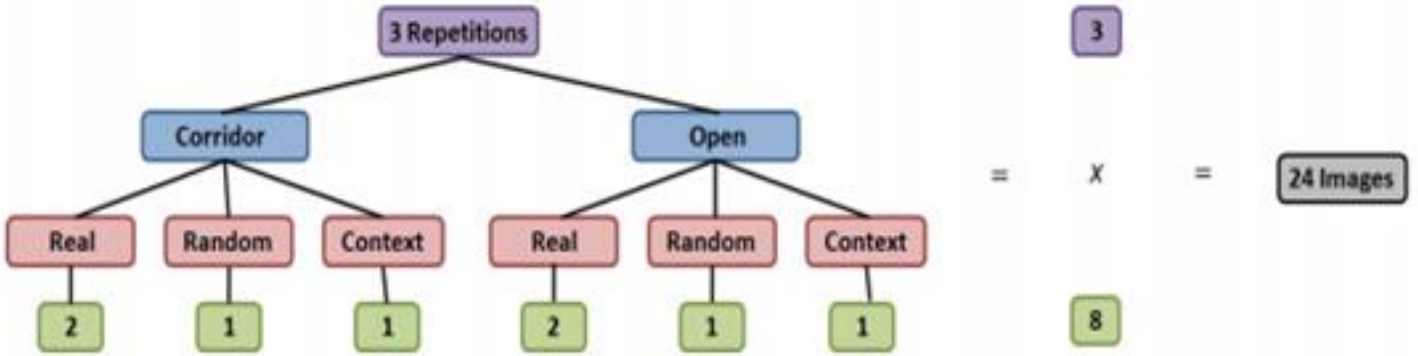
APGV 08
applied perception in graphics and visualization

Experiment

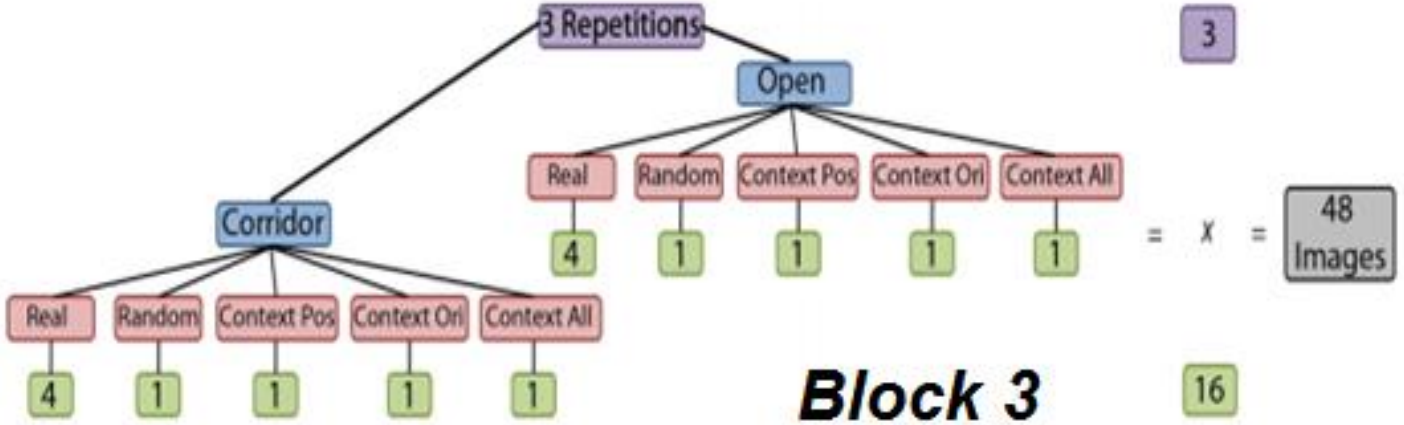
- 32 participants (12F 20M) aged 18 – 30
- 3 Blocks – Position, Orientation, Both
- Participants were asked whether they thought the formation was *Real* or *Synthetically Generated*
- Images displayed for 4 seconds

ROYAL INSTITUTE OF TECHNOLOGY

Experiment 1: Pos and Ori

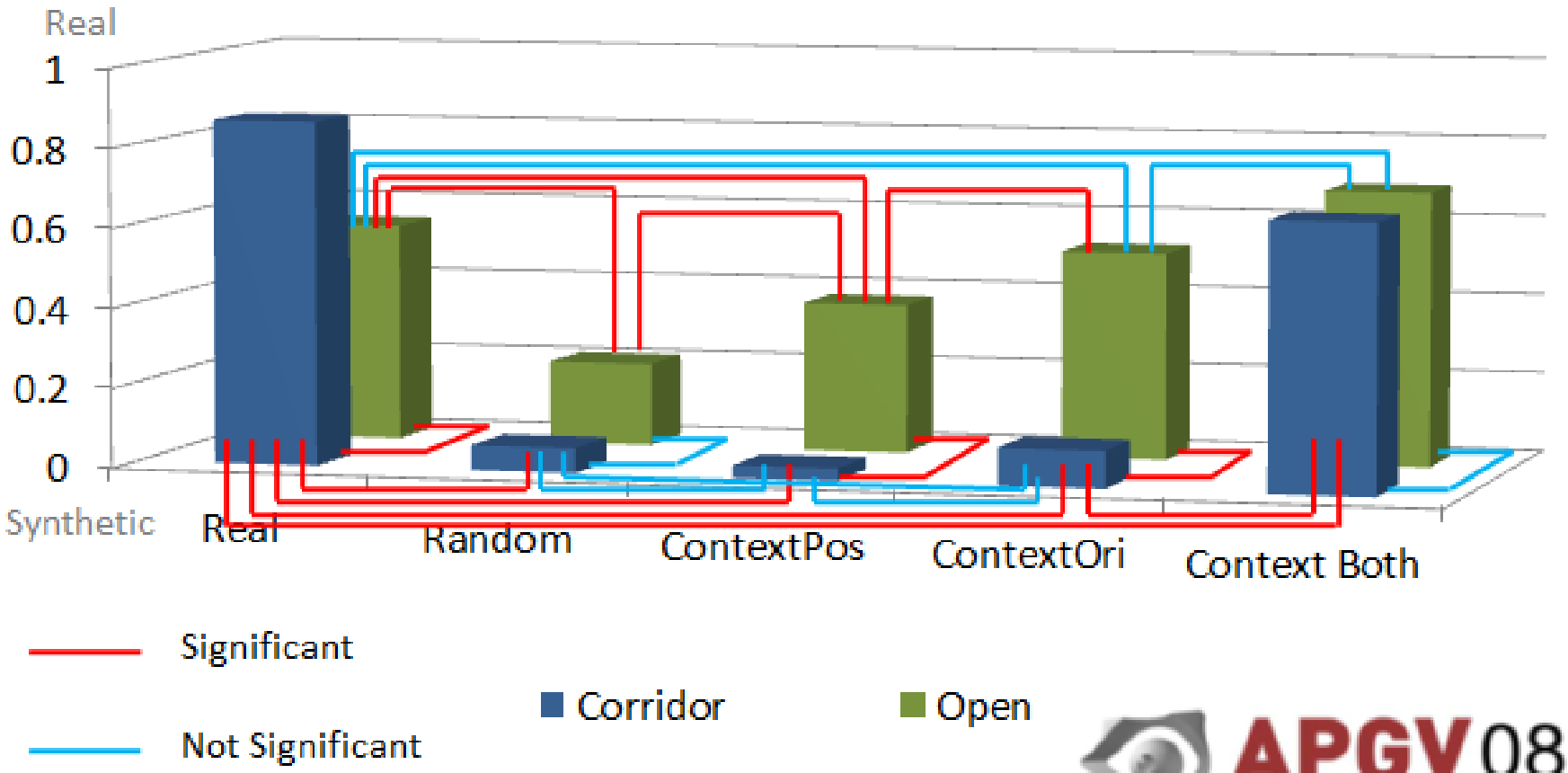


Block 1 and 2



Block 3

Position and Orientation



APGV 08 applied perception in graphics and visualization

ROYAL INSTITUTE OF TECHNOLOGY

Participant Information Sheet

1. Aims and objectives of the study

The aim of the study is to investigate whether participants attend to the gaze direction of a director (in the online communication game). If this is supported, then the main study will investigate the effect of gaze direction (of the director) on the accuracy and response time of participants when selecting objects in common and privileged ground, when instructed to do so by the director.

2. Why have I been chosen?

For the purposes of the study I need to recruit a large number of adult participants who can read and write. This is the only criteria that I have for recruiting people to the study.

3. Do I have to take part?

No. Participation is entirely voluntary. If you change your mind about taking part in the study you can withdraw at any point during the sessions and at any time in the two weeks following that session. You can withdraw by contacting me on email and providing me with your participant information number. If you decide to withdraw all your data will be destroyed and will not be used in the study. There are no consequences to deciding that you no longer wish to participate in the study.

4. What do I have to do?

You will attend one session, lasting approximately one hour. You will complete a series of trials, consisting of the same general stimuli and requiring similar responses. The basic premise of the study is to follow the instructions of the director in moving items around the grid. You will be provided with an example grid as well as two practice grids before the main experiment. You will not be recorded whilst participating in the experiment.

5. What are the risks associated with this project?

The head-mounted display you will wear is designed to be comfortable, but you may feel some initial discomfort. The task itself is very simple.

In case of any discomfort caused by the headsets or by physical exertion caused by the task itself, a researcher trained in First Aid will be present at all times.

6. What are the benefits of taking part?

As an undergraduate student, by taking part in this study you will gain an insight into how a psychology research project is conducted and what it is like to be a participant in such a study. If you are a psychology student this information could be used to shape and inform how you choose to design and conduct your own final year dissertation. You will also gain an insight into the area of virtual agents, online communication paradigms, perspective taking and theory of mind.

As a psychology student, you will receive 60 research participation credits for each session you attend.

7. Withdrawal options

If we have to cancel a testing session I will attempt to contact you as soon as possible using the method indicated by you on the consent form. You will receive research participation credits for any session that has to be cancelled.

If you change your mind about taking part in the study you can withdraw at any point during the session and at any time in the three months following the session by contacting me using the email address stated below. If you decide to withdraw all your data will be destroyed and will not be used in the study.

8. Data protection & confidentiality

The data will be confidential. Only I will have access to the raw data. All the consent forms will be stored in a separate, secure (locked) location from the raw data itself. You will only be identified on the computer records by your participant code number. I will only retain the raw data from the project until final data analyses are completed. They will then be destroyed. When the data has been entered into a computer file, your scores will only be associated with your code number and access to the file will be password protected.

9. What if things go wrong? Who to complain to

If anything goes wrong or you wish to complain about any aspect of the study, please contact [redacted] via email [redacted] or [redacted] explaining the nature of your complaint.

10. What will happen with the results of the study?

The results will be written up for publication in peer reviewed academic journals, as well as potentially at national/international conferences.

11. Who has reviewed this study?

[redacted] has reviewed and passed this study.

12. Further information/Key contact details

Contact [redacted] (via email: [redacted] or phone: [redacted]) for any further details.

ROYAL INSTITUTE OF TECHNOLOGY

Informed Consent Form

Effects of gaze direction of a virtual agent in an online communication game

The aim of the study is to investigate whether participants attend to the gaze direction of a director (in the online communication game). If this is supported, then the main study will investigate the effect of gaze direction (of the director) on the accuracy and response time of participants when selecting objects in common and privileged ground, when instructed to do so by the director.

	Please tick
1. I confirm that I have read and understood the participant information sheet for the above study and have had the opportunity to ask questions.	☐
2. I understand that my participation is voluntary and that I am free to withdraw at anytime without giving a reason.	☐
3. I understand that all the information I provide will be treated in confidence	☐
4. I understand that I also have the right to change my mind about participating in the study for a short period after the study has concluded (*dependent on study start date*).	☐
5. I agree to be recorded as part of the research project	☐
6. I agree to take part in the research project	☐

Name of participant: .....

Signature of participant: .....

Date: .....

Name of Researcher: .....

Signature of researcher: .....

Date: .....

Debrief

The study was designed to investigate the effect of the direction of the gaze of the virtual agent (director) on your response time and accuracy when selecting the items referred to in their instructions. In critical instructions, the item referred to could be one in an occluded slot that was only visible to yourself (that the director had no knowledge of) or one in a mutually visible slot. The latter item was the correct item to select. Previous studies [redacted] have shown that participants do not reliably use the information gained from taking the directors perspective and so select the mutually visible item (i.e. that the director cannot see or has knowledge of the item in the occluded slot); they often select the item in the occluded slot.

In the original task, the directors eye gaze was obscured to constrain this as a variable (pilot studies suggested it randomly interfered with participant item selection). We are aiming to systematically vary where the virtual agent is looking: a between-subjects design was used to implement this.

In Group 1, the eye gaze of the director was focused towards the mutually visible object – this group was predicted to have the fastest response times and highest accuracy due to processing this perspective and therefore cueing selection of the correct item.

In Group 2, the eye gaze of the director was focused towards the occluded object – this group was predicted to have the slowest response times and lowest accuracy due to processing this perspective and therefore cueing selection of the incorrect (competitor) item.

In Group 3, the eye gaze of the director was fixed at the centre of the grid – this group was predicted to provide the baseline of response times and accuracy to which Group 1 is predicted to be faster than, and Group 2 slower than.

This research will provide a starting point in the investigation of the processing of eye gaze of other agents – when, how and what effect this has.

If you have any further questions, please email [redacted] at [redacted] on [redacted]