Exploring novel executive functioning tasks for high-frequency testing using within- and across-task sensitivity

Brónagh McCoy, Lee Casey, Hayley Tseng, Nick Taptiklis, Francesca Cormack

Background

Executive functioning (EF) challenges affect individuals across a vast array of clinical populations, including those living with bipolar disorder, OCD, or ADHD. Several important aspects of executive dysfunction involve inhibition, impulsivity, and working memory (WM). We sought to develop three tasks that would be sensitive to these EF domains, and suitable for smartphone use and high-frequency administration.

Methods

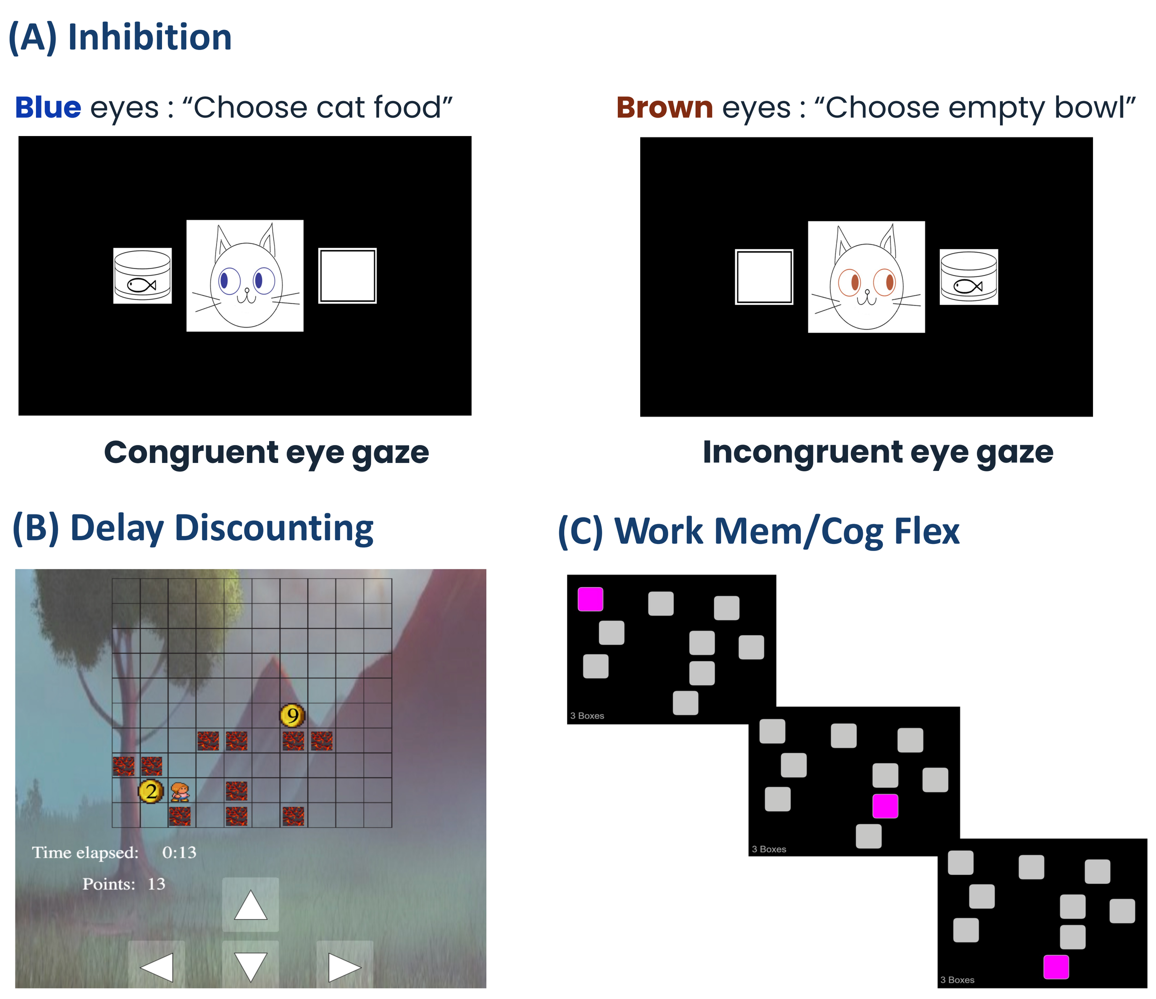
Experiments were created using PsychoPy and jsPsych. Data were collected online via Pavlovia and the Prolific participant recruitment platform as a single battery of three tasks, followed by a brief self-report questionnaire on EF abilities. The tasks ranged from 3–10 minutes in duration. Data were analyzed in R using paired-samples t-tests and linear mixed effects regression modelling. Principal component analysis (PCA) and machine-learning (ML) analyses were carried out using Python and scikit-learn.

Participants:
N=46, age: 39.96 ± 12.59 years, 21 women.

Inhibition / Interference control task:
Participants responded according to a trial instruction indicated by eye colour (see (A) below), while ignoring interference from a task-irrelevant dimension, eye gaze direction. Eye gaze was either congruent or incongruent with the trial instruction. The main outcome measures were 1) response inhibition: RT difference between responding to salient stimulus (cat food) vs. the empty bowl , and 2) interference control: RT difference between congruent and incongruent trials, for the salient (cat food) instruction only.

Delay Discounting task:
Participants chose between two coins of different value by moving an avatar around a grid (see (B) below), with coin values and locations representing larger-later (LL) vs. smaller-sooner (SS) choices. The main outcome measures were proportion of LL choices, proportion of optimal choices, and thinking time before initial movement, with lower scores reflecting more impulsive choices.

Working memory / Cognitive flexibility task:
Participants had to touch squares on a screen (see (C) below) in the temporal order in which they had changed colour (temporal-spatial span (SSP)), or according to the spatial location of the squares that changed colour, either descending from the top or ascending from the bottom of the screen (spatial-SSP versions). The number of squares changing colour, the span length, increased one-by-one when responded to correctly, with a maximum possible score of six.



Measuring across-task sensitivity using PCA and ML:
PCA and ML analyses were carried out on eighteen outcome measures (or combinations thereof) across all three tasks.

EF self-report measures were split into two sub-categories: attentional/WM capacity and impulsivity/hyperactivity symptoms. Median splits of each of these categories provided binary high/low labels for ML analyses.

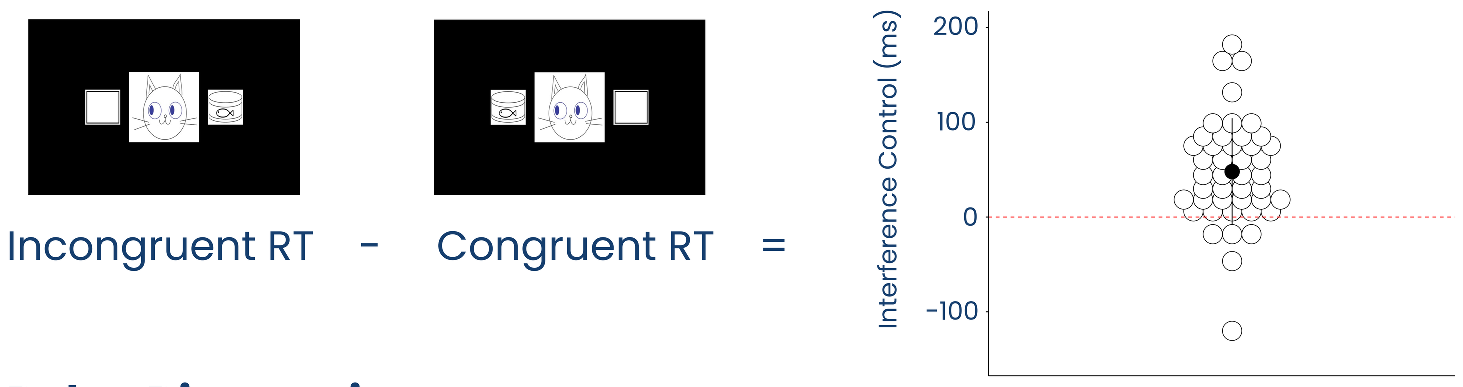
Six ML classifiers – logistic regression, support vector machines, decision trees, random forest, naïve bayes, and k-nearest neighbours, were applied and compared using a 3-fold cross-validation procedure.

References

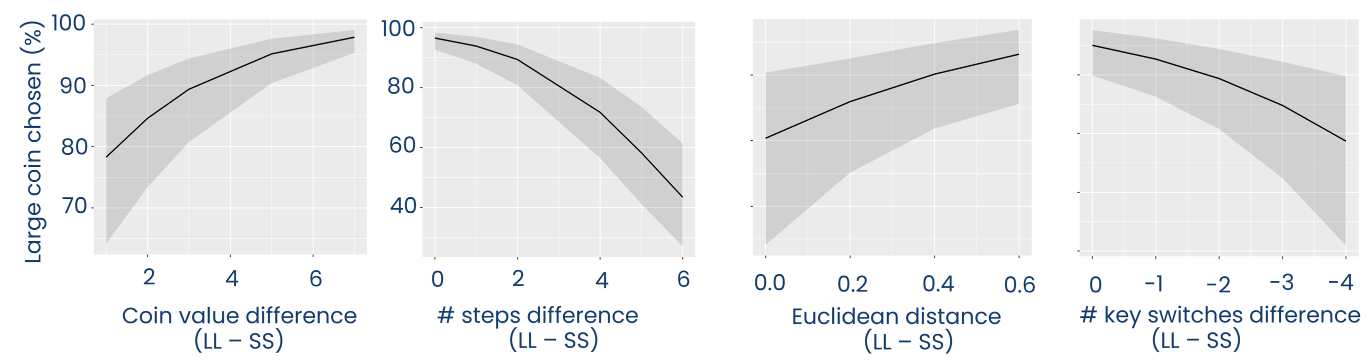
Diamond, Annu Rev Psychol, 2013
Bialystok et al., J Exp Psychol Learn Mem Cogn, 2006
Scherbaum, PLOS One, 2013
Scherbaum, J Clin & Exp Neuropsych, 2018

Results

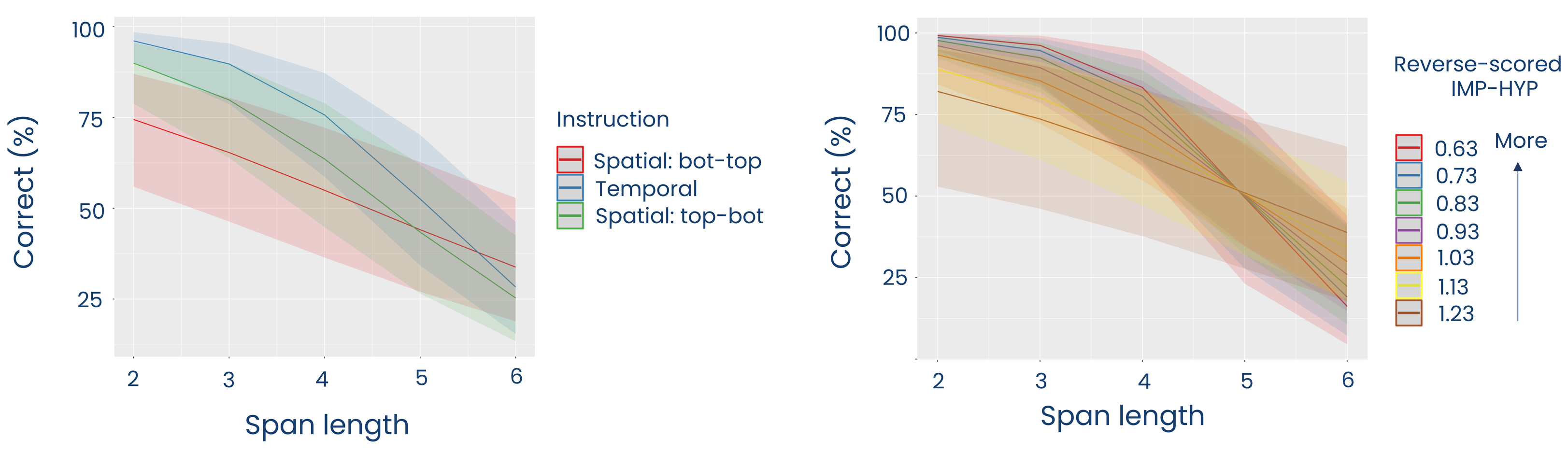
Inhibition / Interference control:
There was a robust interference control effect, captured by an RT increase in "choose cat food" trials when there was interference from a task-irrelevant distraction (47.96 ± 56.31 ms; t(43)=5.65, p<.001), i.e. when eye gaze was incongruent vs. congruent. There was no relationship found between these outcome measures and EF sub-domains.



Delay Discounting :
For each of the main outcome measures - proportion of LL choices, proportion of optimal choices, and thinking time – there were strong positive effects of LL vs. SS coin value difference (e.g. proportion LL: OR =1.53, p<.001), number of additional direction changes to reach the LL vs. SS coin (e.g. proportion LL: OR =0.66, p<.001), and number of steps to either the LL or SS coin, or their difference (e.g. proportion LL: OR =0.55, p<.001), validating that participants were overall attentive to specific trial conditions when making their choices. There was no relationship found between any of these outcome measures and EF sub-domains.



Working memory / Cognitive flexibility:
Participants performed worse with increasing span length (OR=0.65, p<.001), better overall in the temporal-SSP block than in the spatial adaptations (e.g. temporal vs. ascending: OR=27.84, p<.001), and better in the descending compared to ascending spatial-SSP version (OR=6.66, p<.001). There was a significant relationship between performance in both temporal-SSP (shown below) and descending spatial-SSP and measures of impulsive/hyperactive (IMP-HYP) symptoms (p=.028 and p=.009, respectively). Participants with fewer IMP-HYP symptoms performed worse in general, and worse at lower span lengths but better for increasing span length than those with more IMP-HYP symptoms (span * IMP-HYP score: p=.006 and p=.007, for temporal- and spatial-SSP, respectively).



Across-task sensitivity :
PCA across all task outcomes resulted in three PCs. The most influential PC was driven by a negative weighting on the interference control effect, and several difference measures from the temporal-SSP task. Weightings were obtained for each PC per participant and tested for a relationship with self-report ATT-WM or IMP-HYP EF measures. There were no significant correlations found between these measures (all p>.1).

The best ML model for predicting self-reported high vs. low ATT-WM abilities used a support vector machine (SVM) classifier (acc: 81.62%, prec: 91.67%, recall: 80.56%, AUC: 83.61%). This had three task outcomes measures as inputs: the interference control effect in the inhibition task, RT instruction difference for congruent trials only in the inhibition task, and % optimal choices in the delay discounting task. No other data (e,.g. demographic) was used to train this classifier.

Conclusion

Combined use of several tasks to capture distinct and overlapping dimensions of EF is a useful tool that can be applied to clinical populations. Here we tested short EF tasks that would be suitable for smartphone use and validated several robust outcome measures for each task. Although PCs based on all across-task outcome measures did not show any reliable correlation with self-reported EF domains, a ML approach using SVM could predict participants with low or high ATT-WM EF abilities with high accuracy. EF questionnaire measures that are more sensitive and specific than the brief questionnaire used here will likely improve and refine across-task sensitivity to these domains.

Authors & contacts

Bronagh.Mccoy@camcog.com, R&D Scientist, CC
Nick.Taptiklis@camcog.com, Head of R&D Technology, CC
Francesca.Cormack@camcog.com, Chief Scientist, CC

CC: Cambridge Cognition Ltd., Tunbridge Court, Tunbridge Lane, Bottisham, Cambridge, CB25 9TU, UK
BM, NT, FC are current employees of CC; LC and HT are former interns.
The authors report no conflicts of interest for this work.

CAMBRIDGE COGNITION