

How Does AI Moderation Influence Decision-Makers' Information Search And Accuracy?



Eoin Cremen¹, Dr Janina Hoffmann¹, Prof Konstantinos Katsikopoulos²

ec531@bath.ac.uk

¹ University of Bath, ²University of Southampton



Summary

Individuals often search for additional information to make accurate decisions.

In the example of a physician, several diagnoses may be considered and tests are conducted to identify the most likely one.

AI is increasingly valuable as a source of information and advice.

We designed three AI moderators: **Devil's Advocate**, **Moderator**, **Facilitator**.

We found that participants adapted search strategies following advice.

Research Questions & Variables

RQ 1: How do the information search strategies of decision-makers respond to advice delivered by an AI moderator?

RQ 2: How does moderating AI advice influence decision-makers' accuracy?

Diversity of first test selections: Gini-Simpson Index value of which tests were selected first by each participant.

Accuracy: Binary measure of correct or incorrect diagnosis in each trial.

$$GSI = 1 - \sum_{i=1}^s p_i^2$$

Methods

Within-subjects, repeated measures design.

85 psychology undergraduates played a medical information search game.

Learning phase: Feedback after each trial enabled participants to learn the relationships between the disease, symptom, and test results.

Test Phase 1: Participants chose medical tests to reveal before diagnosing the patient.

Participants were not given feedback on accuracy.

Test selection strategies in Phase 1 were used to inform the details of the advice in Phase 2.

Test Phase 2: AI moderators provided text advice before first test selection.

A control condition (no advice) was included during phase 2.

The Task

Cue: view 1 of 4



Advice: 4 conditions randomised

Tests: select up to 4



Disease: select 1 of 4

Results: 3 outcomes possible per test



Score: \$1000 per correct diagnosis



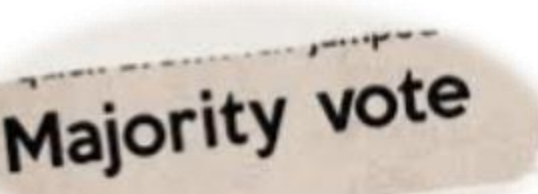
Devil's Advocate encouraged counterfactual information.

"consider a DIFFERENT test this time."



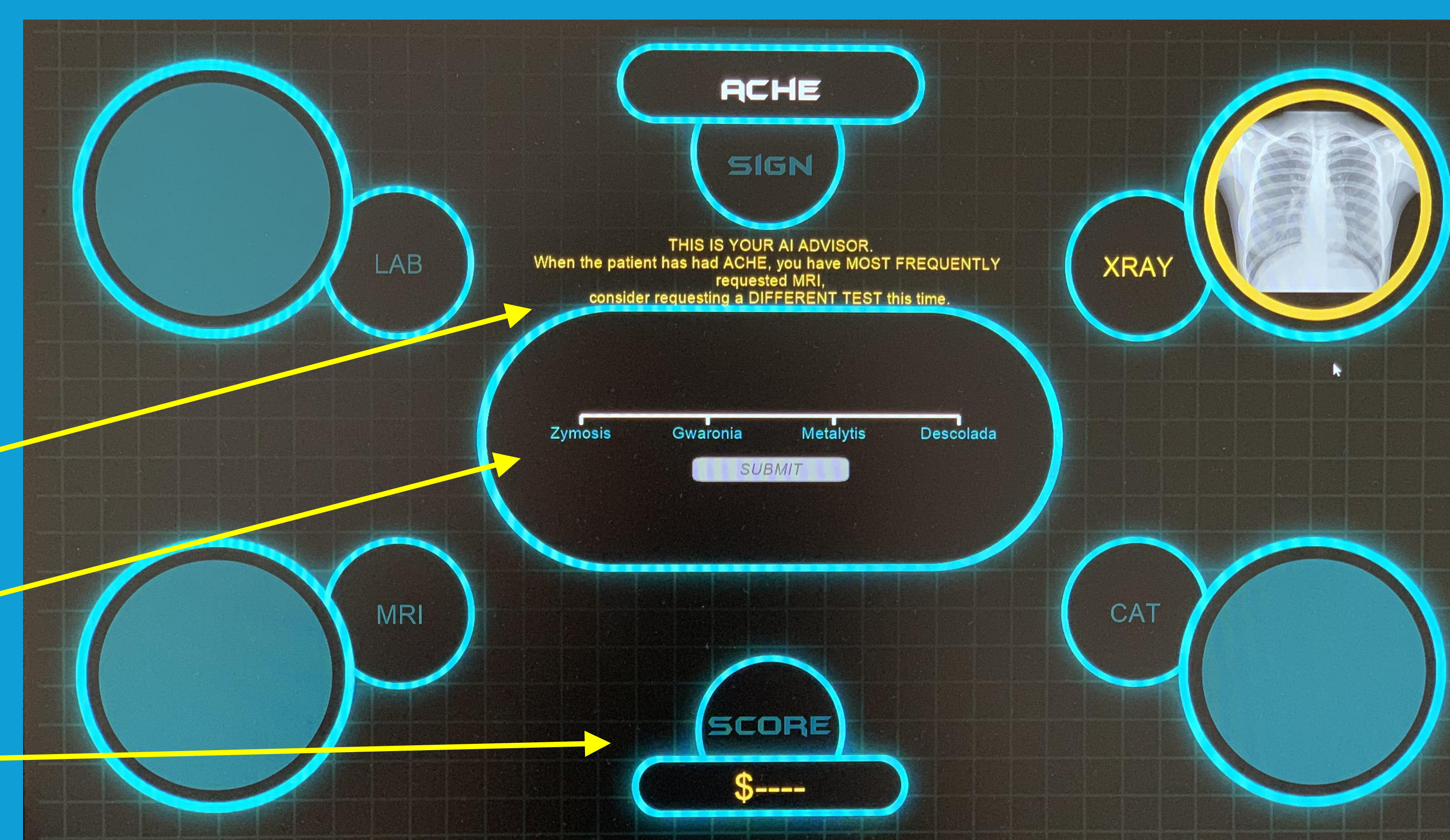
Moderator encouraged least preferred information.

"consider LEAST FREQUENT TEST this time."



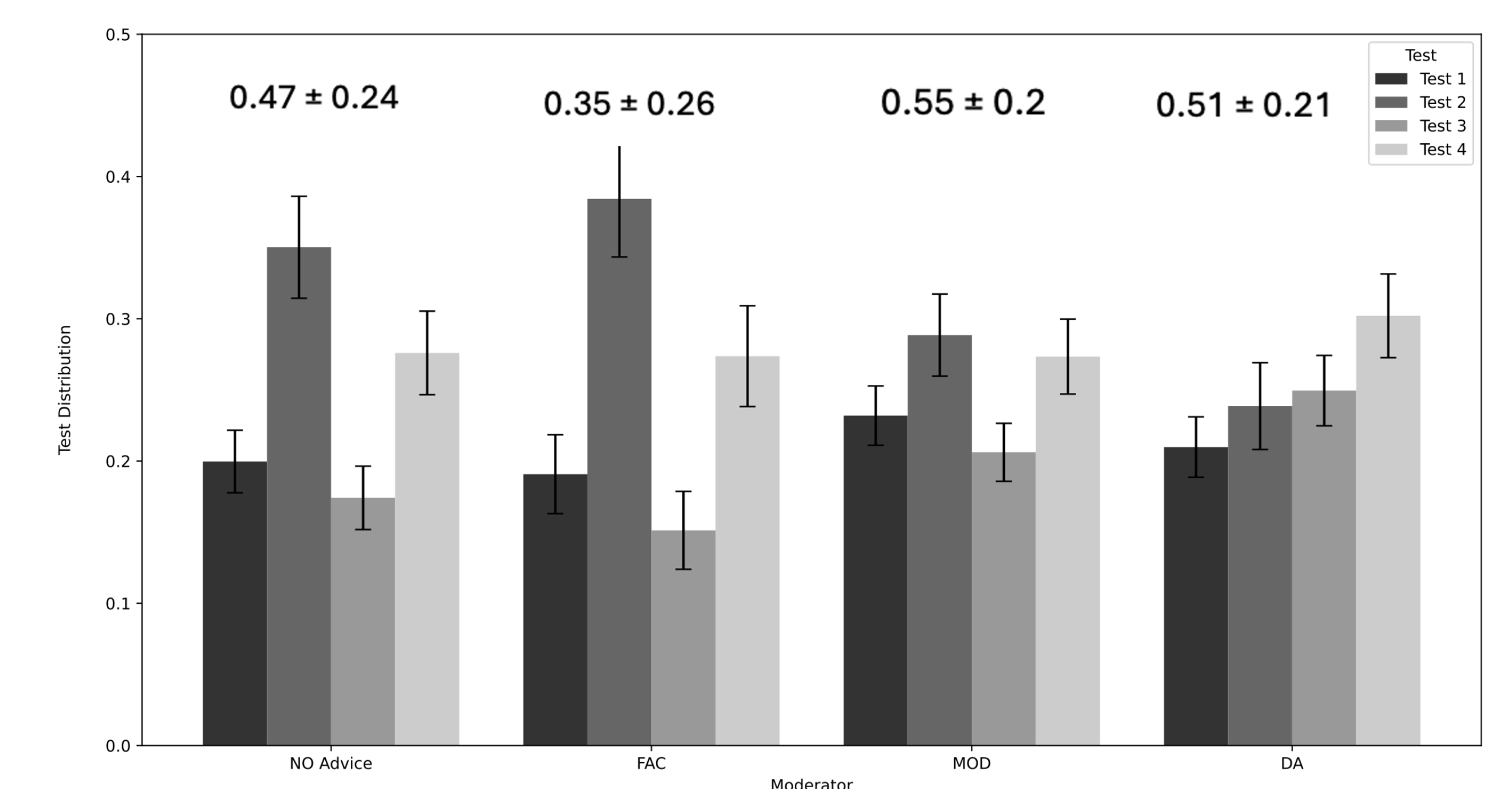
Facilitator encouraged most preferred information.

"consider THIS MOST FREQUENT TEST again."



Results

Diversity of Test Selections (GSI ± SD)

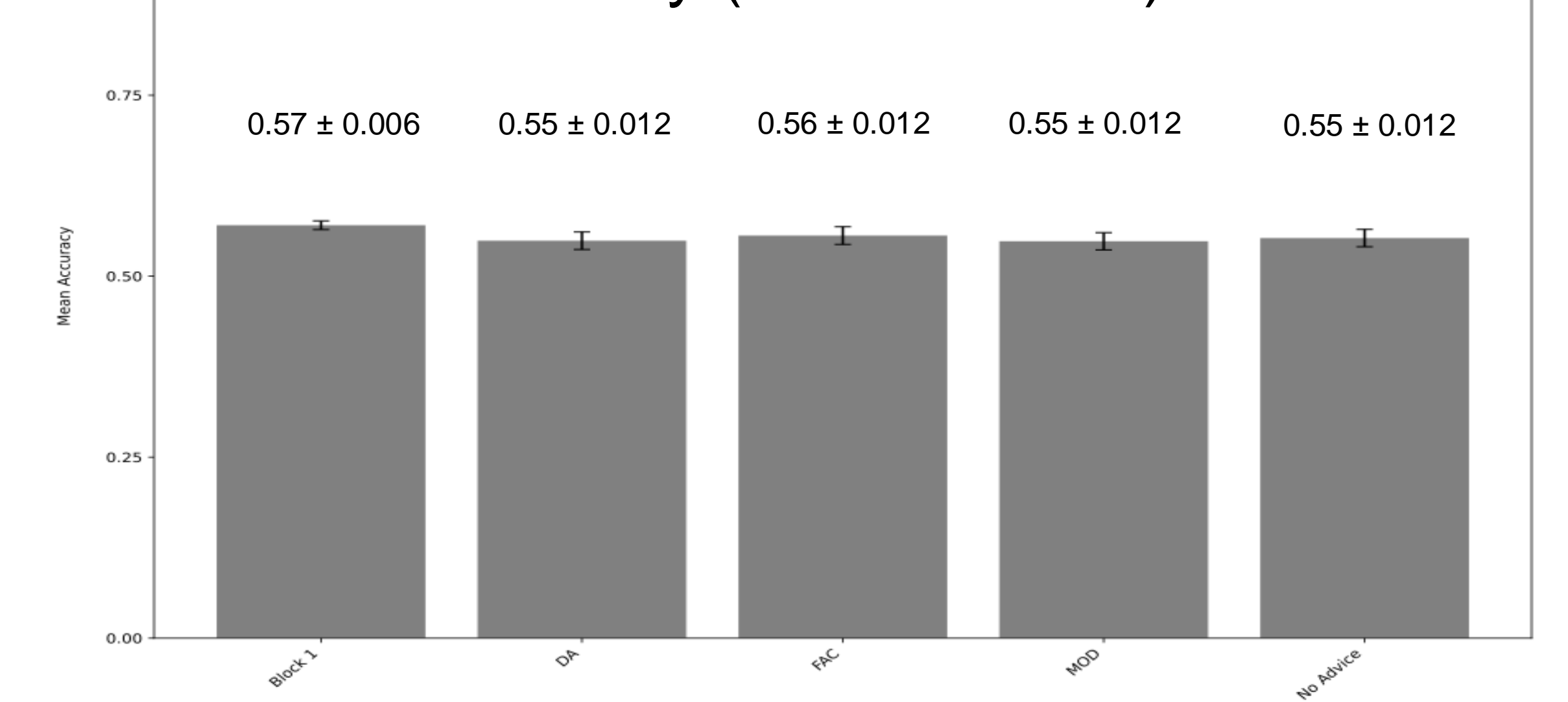


FAC reduced diversity and terminated information search earlier.

DA + MOD increased diversity and extended information search.

Accuracy was not influenced by any condition.

Accuracy (Mean ± SEM)



References

Cialdini, R.B. and Goldstein, N.J., 2004. Social Influence: Compliance and Conformity. Annual Review of Psychology, 55(1), pp.591–621.

De Dreu, C.K.W. and West, M.A., 2001. Minority dissent and team innovation: The importance of participation in decision making. Journal of Applied Psychology, 86(6), pp.1191–1201.

Illingworth, D.A. and Thomas, R.P., 2022. Strength of Belief Guides Information Foraging. Psychological Science.

Nemeth, C., Brown, K. and Rogers, J., 2001. Devil's advocate versus authentic dissent: stimulating quantity and quality. European Journal of Social Psychology 31(6), pp.707–720.

Xiao, E., 2017. Justification and conformity. Journal of Economic Behavior & Organization, 136, pp.15–28.