Overview of the Flanker Attention Task (Left or Right)

Authors: Roger Strong and Amy Price

Table of Contents:

Background & Scientific Purpose

The ability to focus on relevant information while filtering out distractions is crucial for everyday functioning. This skill is referred to as cognitive control, and it is essential in a range of life situations, from reading in a noisy coffee shop to navigating a busy road.

Classic flanker paradigms are designed to measure cognitive control and have been used in research studies for decades (Eriksen & Eriksen, 1974). They have been especially useful as an assessment in contexts where a reduced ability to filter out irrelevant information is suspected, such as in childhood ADHD (Mullane et al., 2009). The TestMyBrain Flanker Attention task (Erb et al., 2023; Treviño et al., 2021; 2022) is an adapted version of the classic flanker paradigm and is designed for remote administration. The TestMyBrain Flanker Attention task has previously been used to investigate cognition both in cancer survivors (Treviño et al., 2022) and control samples (Erb et al., 2022; Treviño et al., 2023).

Methodology

On each trial of the TestMyBrain Flanker Attention task, participants are instructed to report the direction (left or right) of a target arrow presented at fixation, which is flanked on each side (left and right) by two arrows. On half of the trials, the target arrow points in the same direction as the flanker arrows (congruent trials), while on the other half of trials the target arrow points in the opposite direction of the flanker arrows (incongruent trials) (see Figure 1). The test measures

the extent to which each participant's response time and accuracy worsens for incongruent trials as compared to congruent trials. The less the difference in performance between congruent and incongruent trials, the better the participant's cognitive control.

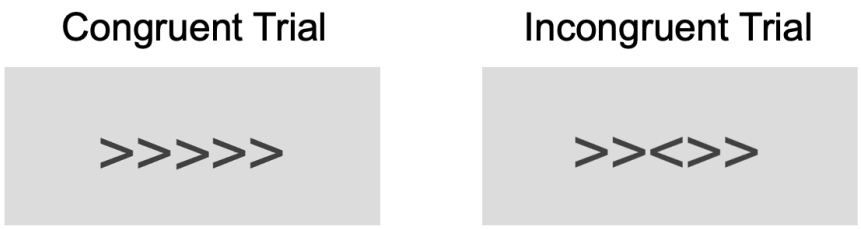

Figure 1: Example of arrow arrangements during congruent and incongruent trials.

Figure 2 provides an overview of the test's structure. At the start of the test, the participant's response input method is determined. Participants without touch-compatible devices are assigned to using keyboard input, whereas participants using devices with touch input must select whether they will use touch or keyboard input.

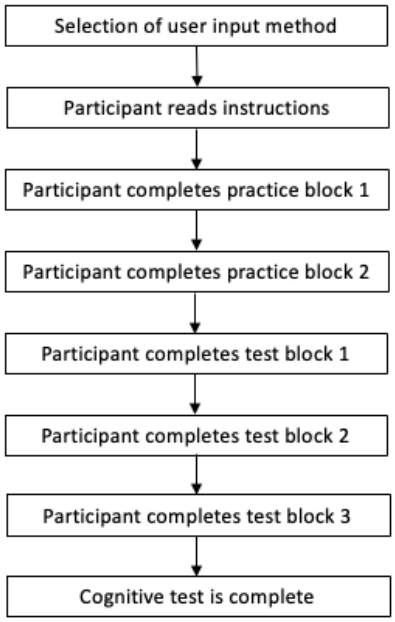

Figure 2: Overview of Flanker Attention test structure

After the response selection method is determined, participants view brief instructions for the test (Figure 3), then complete the first of two rounds of practice trials. For the first practice round, participants complete four trials and the trials move at a slower pace than the test trials.

Each trial begins with only the response option buttons present (Figure 4). After 1000 ms, the flanking arrows appear, all pointing in the same direction (left or right), without the central target arrow. After 600 ms, the target arrow appears along with the flanking arrows for 300 ms. Finally, all the arrows are replaced by a fixation cross until the participant makes a response. For the practice trials, if a participant answers incorrectly or does not respond within three seconds of the flanker arrows appearing, a message appears warning the participant that they did not make the correct selection. When this occurs, participants must repeat the trial.

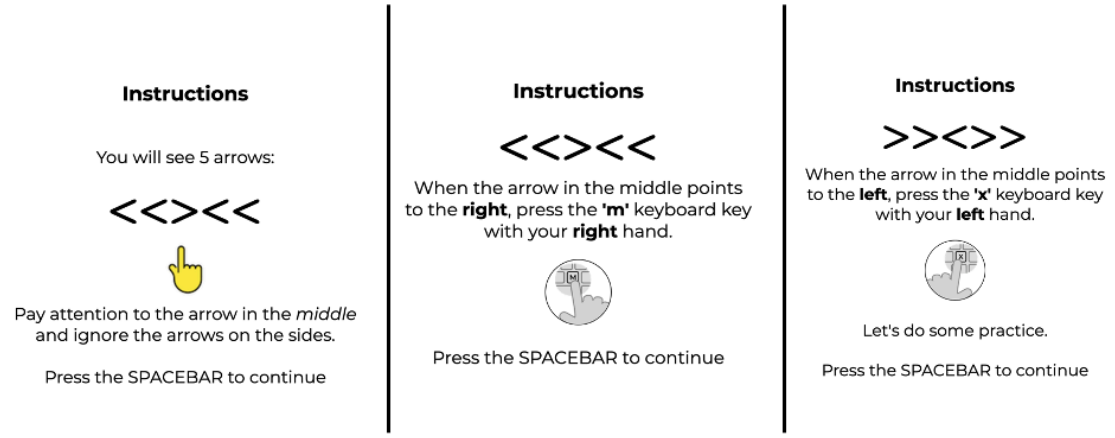

Figure 3: Test instructions for participants using keyboard input.

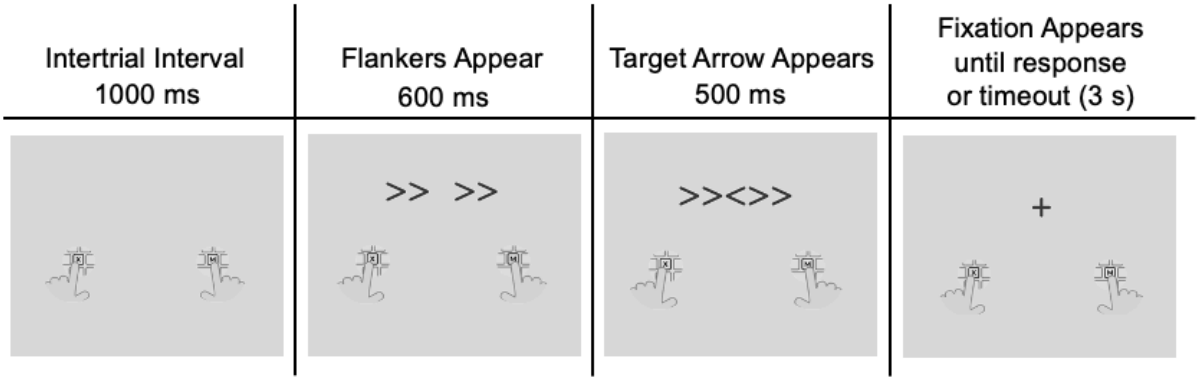

Figure 4: Structure of first round of practice trials (keyboard input)

Participants next view a brief reminder of the task's instructions, and then complete the second and final round of practice trials. The second round of practice trials has the same timing and structure of the upcoming test trials (Figure 5). Participants again complete four trials, with each trial beginning with only the response option buttons present. After 1000 ms, four flanking arrows appear, all pointing in the same direction (left or right), without the central target arrow. After 100 ms, the target arrow appears along with the flanking arrows for 300 ms. After the target arrow has been visible for 300 ms, all the arrows are removed from the screen, leaving

only the response button options visible. After 70 ms, a fixation cross appears and remains visible until the participant makes a response, or the trial times out after 3 seconds. Like the first block of practice trials, if a participant answers incorrectly or does not respond within 3 seconds, a message appears warning the participant that they did not make the correct selection. When this occurs, participants must repeat the trial.

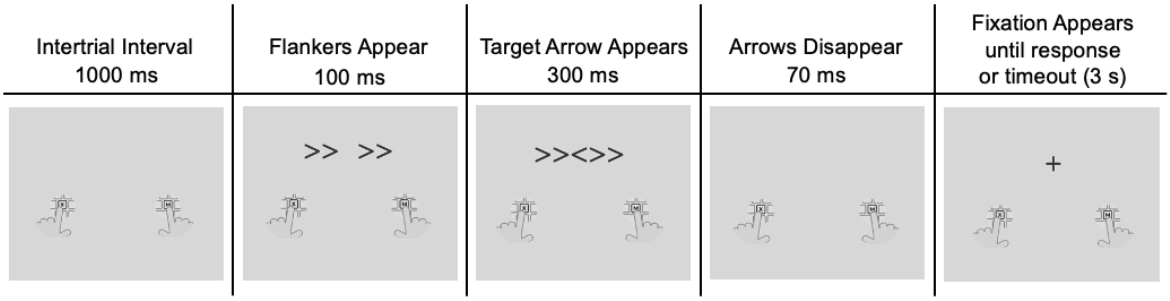

Figure 5: Structure of second round of practice trials and test trials.

After the second round of practice trials is completed, participants view a brief reminder of the task instructions before completing three rounds of test trials, each of which contains 32 test trials (96 total test trials). Each round of test trials has an equal number of trials where the target arrow points left versus right, and an equal number of congruent versus incongruent trials. During the test trials, participants are not informed when they make an incorrect response.The participant only repeats test trials if the trial times out, which occurs if they fail to make a response within three seconds of the flanker arrows being presented.

Between each of the three test trial rounds, participants view a message updating them on their progress and providing a reminder to respond as quickly and accurately as possible. If the participant made four or more incorrect responses during the prior test trial round, then the progress message emphasizes responding accurately. If the participant made two or fewer errors during the prior test trial round, then the message emphasizes responding quickly. If the participant made exactly three incorrect responses during the prior test trial round, no emphasis is placed on either speed or accuracy.

Data & Analysis Guidelines

Data

As described in the Introduction to Cognitive Testing Data in the All of Us Research Program *Support Hub Article*, there are three main categories of data available for cognitive tests: (1) trial-level data, (2) summary scores, and (3) metadata. Please see the Exploring the Mind Data

Dictionary for a description of the trial-level data (trial_data), summary score (outcomes), and metadata (metadata) variables for this test (Flanker).

Suggested Outcomes

The test's suggested primary outcome is rate correct score interference (*rcs_interference*). For each participant, a rate-correct score can be calculated separately for congruent and incongruent trials; the rate-correct score is computed as: proportion of trials answered correctly divided by median reaction time in seconds. Rate correct score interference is the difference in rate correct score for congruent trials and rate correct score for incongruent trials. However, researchers may also consider using reaction time interference (*medianRT_interference*) or accuracy interference (*accuracy_interference*), or computing a different aggregated measure of reaction time interference and accuracy interference.

Outcome Type	Outcome Name	Description
Primary	rcs_interference	Difference in rate correct score between congruent and incongruent trials. Larger values indicate greater interference due to incongruent arrows.
Secondary	medianRT_interference	Difference in median reaction time between incongruent and congruent trials. Larger values indicate greater interference due to incongruent arrows.
	accuracy_interference	Difference in accuracy between congruent and incongruent trials. Larger values indicate greater interference due to incongruent arrows.

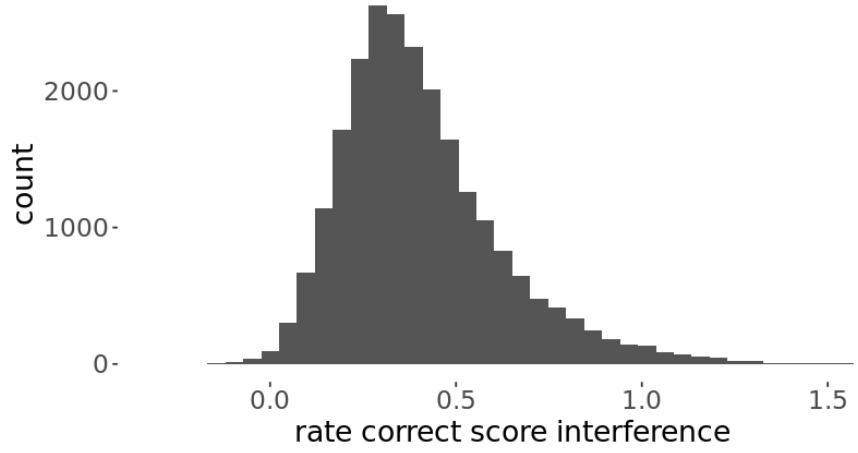

Figure 6: Histogram of primary outcome metric (rate correct score interference), for all participants in the CDR v8 off-cycle release

Quality Control Guidelines

The following guidelines are provided for the purpose of flagging extreme deviations in performance from what is typically seen in participants performing the task in a valid manner. Researchers must use their own judgment when determining whether flagged participants should be excluded from analyses. Researchers may also consider implementing their own quality control criteria separately from these recommendations. For more details about quality control criteria, please see Introduction to Cognitive Testing Data in the All of Us Research Program.

Quality control variables are provided both in trial-level data and full-test outcomes data. The table below summarizes the quality control variables available for this test.

Flag Type	Variable Name	Description
Trial-level	flagged	Indicates whether a participant's reaction time (rt) for a test trial is less than 200 ms (1 if rt < 200, 0 if rt >= 200, null if no response and for practice trials). Accurately responding in under 200 ms on this task is physiologically implausible.
Full-test	flag_medianRT	Has a value of 1 if the participant has a median reaction time under 400 ms (medianRT < 400), and a value of 0 otherwise. A median reaction time under 400 ms is implausibly fast for this task and suggests the participant was responding carelessly.
	flag_accuracy	Has a value of 1 if the participant answers fewer than 65% of trials correctly (accuracy < .65), and a value of 0 otherwise. Because chance-level performance on the task is 50% accuracy, answering fewer than 65% of trials correctly suggests that the participant was randomly guessing on most trials.
	flag_trialFlags	Has a value of 1 when more than 10% of a participant's trials are flagged for having reaction time (rt) less than 200 ms (flagged=1), and a value of 0 otherwise. The accuracy of prior participants completing this test was distributed around chance-level (.50) when more than 10% of trials were answered in less than 200 ms, suggesting these participants were randomly guessing to complete the test quickly.

Flanker Task (N = 23,727)[1]		
	Yes	**No**
Median RT Flags	<1%[2]	>99%
Accuracy Flags	1.2%	98.8%
Trial Flags	<1%	>99%
Any Flags	1.2%	98.8%

Table 1: Percentage of participants with quality control flags in the Exploring the Mind CDR v8 off-cycle release.

Calculating Test Reliability

To calculate the reliability of Flanker performance differences between participants in a given sample, we recommend calculating *split-half reliability* (Pronk et al., 2022) using the following steps:

- For each participant, separate congruent trials and incongruent trials.
- Mark whether the sequential order of each trial was “odd” or “even.” For example, the first congruent trial would be marked “odd” and the second congruent trial would be marked “even.” Correspondingly, the first incongruent trial would be marked “odd” and the second incongruent trial would be marked “even.”
- Separately for odd and even trials, compute rate-correct score for congruent trials (proportion of trials answered correctly divided by median reaction time in seconds)
- Separately for odd and even trials, compute *rate-correct score* for incongruent trials (proportion of trials answered correctly divided by median reaction time in seconds)
- Separately for odd and even trials, compute *rate-correct score interference* (rate-correct score for congruent trials minus rate-correct score for incongruent trials).
- Compute the Pearson correlation (r) between (1) *rate-correct score interference* on odd trials and (2) *rate-correct score interference* on even trials.
- Use the Spearman-Brown prediction formula to compute full-test reliability: reliability = (2*r) / (1+r)

[1]This count is defined as the total number of unique participants who completed the task.
[2]Due to the data dissemination policy, counts of less than 20 participants cannot be shared publicly. Users can view exact counts in the corresponding featured workspace after logging into their Researcher Workbench account.

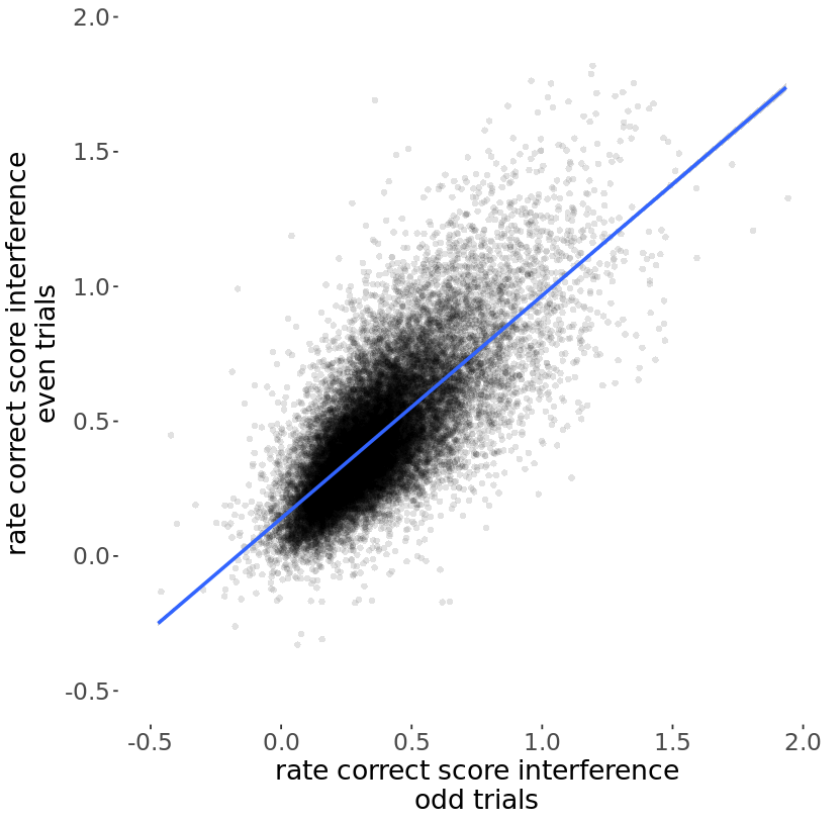

Figure 7: Correlation of participants' Flanker rate correct score interference on even and odd trials (Spearman-Brown split-half reliability = .86)

Correlates of Interest

Prior data collection has found associations between the following demographic variables and performance (*accuracy*) on TestMyBrain Flanker Attention. Therefore, researchers may consider including the following variables as covariates in analyses.

1. *age:* increasing age has been associated with larger Flanker reaction time interference and accuracy interference (Erb et al., 2023). Reaction time interference increased consistently from younger to older ages, whereas accuracy interference decreased from younger ages until around age 40, after which it started increasing with older age.
2. r*esponse input format*:
    a. In *All of Us* Exploring the Mind and TestMyBrain.org validation data, on average the rate correct score interference was larger for participants using keyboard input than for participants using touch input.

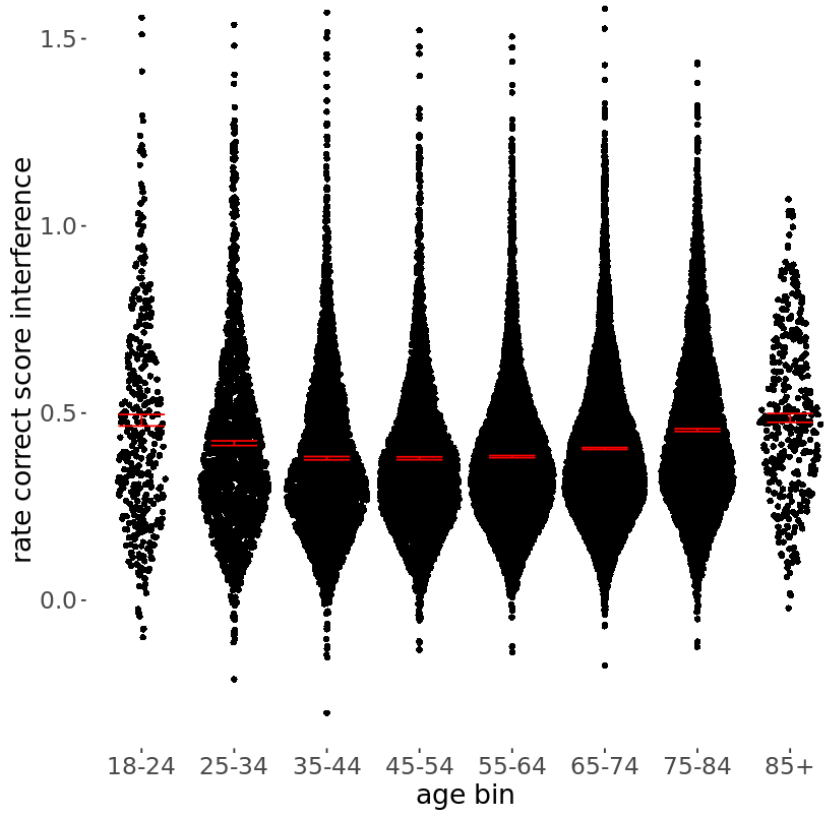

Figure 8: Accuracy interference by age bucket. Red lines represent mean accuracy interference for each bucket. Width of distributions (black dots) represent the relative density of participants at each magnitude of accuracy interference.

References

Erb, C. D., Germine, L., & Hartshorne, J. K. (2023). Cognitive control across the lifespan: Congruency effects reveal divergent developmental trajectories. *Journal of Experimental Psychology: General, 152*(11), 3285–3291. https://doi.org/10.1037/xge0001429

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143-149. https://doi.org/10.3758/BF03203267

Mullane, J. C., Corkum, P. V., Klein, R. M., & McLaughlin, E. (2009). Interference control in children with and without ADHD: a systematic review of Flanker and Simon task performance. *Child Neuropsychology*, *15*(4), 321-342. https://doi.org/10.1080/09297040802348028

Pronk, T., Molenaar, D., Wiers, R. W., & Murre, J. (2022). Methods to split cognitive task data for

estimating split-half reliability: A comprehensive review and systematic assessment. *Psychonomic Bulletin & Review, 29*(1), 44-54. https://doi.org/10.3758/s13423-021-01948-3

Treviño, M., Zhu, X., Lu, Y. Y., Scheuer, L. S., Passell, E., Huang, G. C., ... & Horowitz, T. S. (2021). How do we measure attention? Using factor analysis to establish construct validity of neuropsychological tests. *Cognitive Research: Principles and Implications, 6*(1), 1-26. https://doi.org/10.1186/s41235-021-00313-1

Treviño, M., Zhu, X., Lu, Y. Y., Scheuer, L. S., Passell, E., Huang, G., ... & Horowitz, T. S. (2022). Establishing construct validity of neuropsychological tests in cancer survivors. *Psycho‐Oncology, 31*(10), 1728-1736. https://onlinelibrary.wiley.com/doi/abs/10.1002/pon.6015