Introduction

Over the last decades, life expectancy has steadily increased and is predicted to further increase in the coming years (Kanasi et al., 2016). Although age-related changes in cognitive functions, such as executive control, attention, and memory, have been repeatedly demonstrated (for a review see: (Rey-Mermet and Gade, 2018; Verhaeghen and Cerella, 2002), the underlying processes remain largely unknown.

An executive function that is particularly affected by aging is inhibitory control, the ability to suppress highly practised responses in favor of more appropriate reactions given the current context or goals (Butler and Zacks, 2006; Connelly et al., 1991; Crawford et al., 2005; Houx et al., 1993; Rey-Mermet and Meier, 2017; Spieler et al., 1996). Recently, the voluntary control of eye movement has been proposed as a simple to use, non-invasive, and potentially clinically relevant method to measure inhibitory control using the antisaccade task (Antoniades et al., 2013; Crawford et al., 2017, 2005; Shafiq-Antonacci et al., 2003). In the antisaccade task, participants are instructed to suppress a reactive eye movement (prosaccade) to a sudden onset of a laterally presented visual stimulus, in order to execute a voluntary eye movement (antisaccade) to a point in the visual field opposite the target (Hallett, 1978; Ramat et al., 2007). It is generally assumed (e.g. Peltsch et al., 2011) that reduced ability to inhibit the prepotent saccade typically results in slower responses or higher incorrectness in the antisaccade task (Butler and Zacks, 2006; Sweeney et al., 2001), which has been repeatedly found in elderly participants as compared to younger controls (Abel and Douglas, 2007; Bojko et al., 2004; Klein et al., 2000; Sweeney et al., 2001). However, these studies mainly focussed on average reaction times and error rates when evaluating participant's task performance and overlooked different sources of a worse performance of erderly participants as compared to younger

controls during the antisaccade task (Reuter et al., 2005; Wiecki and Frank, 2013). Therefore, we aim to report full reaction time and error rate distributions (Cutsuridis et al., 2007; Feng, 2012; Meeter et al., 2010) and additional measures, like peak saccadic velocity and the saccade gain, as proposed in the internationally standardized antisaccade protocol (Antoniades et al., 2013).

Additionally, we use a probabilistic computational model to study the antisaccade task, referred to as the Stochastic Early Reaction, Inhibition and Late Action Model (SERIA, Aponte et al., 2017), which links the concept of competing early processes (Camalier et al., 2007; Logan et al., 1984) with two voluntary actions that generate late pro- and antisaccades. This formal probabilistic approach will enable us to analyze the metrics not detectable by error rates and reaction time measures, especially inhibition failures, which are fast, reflexive prosaccades, which would be correct on prosaccade trials and errors on antisaccade trials (Aponte et al., 2019).

Moreover, previous studies typically recorded antisaccade task performance at a single time point per participant (Abel and Douglas, 2007; Peltsch et al., 2011) and thus it remains unknown whether antisaccade task metrics provide reliable estimates of inhibitory control for individual subjects - a prerequisite in order to qualify for clinically relevant markers of cognitive impairment. In order to bridge this gap, we further evaluate the test-retest reliability across two testing sessions per participant one week apart. In reference to our design analysis (reported in the Methods section), a total of 156 healthy participants (based on our power analysis) from two age groups (i.e.,78 young adults: age range: 20-35 years; 78 elderly adults: age range: 60-80 years) will take part in a test-retest experimental design.

Based on the literature and our pilot study (see section: Pilot Data), we hypothesize:

1.Significantly higher average error rates for older as compared to younger adults in the antisaccade task.

2. Longer saccadic reaction times for older adults as compared to younger adults in the antisaccade task.

3. High test-retest reliability (for reaction times, peak saccade velocity and gain indicating excellent or good reliability, i.e., intraclass correlation coefficient (ICC) > 0.6; McGraw and Wong, 1996).

4. Based on the SERIA model by Aponte (2018), we expect significantly more inhibition failures for older adults as compared to young adults. Inhibition failures are classified as fast, reflexive prosaccades on prossacade trials and errors on antisaccade trials.

Materials and Methods

Dataset Description

The data to be used in this study is currently being recorded in our laboratory in the context of a larger project that aims to quantify age-effects on eye movement behavior and electroencephalography (EEG) recordings of resting-state and task-related activity. A total of 200 subjects (the first 44 subjects recorded between January and April 2019 are considered pilot subjects (see Pilot Data section), the remaining 156 subjects recorded from May 2019 onward, will be used for the main analysis and these data have not been observed before), from two age groups (i.e. 100 young adults: age range: 20-35 years; 100 elderly adults: age range 60-80). will take part in a test-retest experimental design, in which the same data recordings are performed one week apart (at the same time of day). Each recording includes a test battery of seven experimental paradigms assessing key cognitive functions affected by age, such as visual perception, attention, working memory, episodic memory, cognitive control, and processing speed (Kozak and Cuthbert, 2016). For the purpose of the present study, we focus on the eye-tracking data from the antisaccade task. The study was approved by the local ethics commission. All participants gave their written informed consent prior to participation in the study and received a monetary compensation (the local currency equivalent of USD 25).

For exploratory analysis, hypothesis generation and technical validation of our data processing pipeline, we conducted an analysis of a pilot dataset (described in the "Pilot Data" section). To further increase the transparency of our planned analyses, all processing scripts and data collected from our ongoing study can be found online in an OSF repository.

Power Analysis

In order to estimate the sample size needed in our study, we performed a literature search and found 10 studies that compared antisaccade task performance between young and elderly adults (Bialystok et al., 2006; Bojko et al., 2004; Butler et al., 1999; Butler and Zacks, 2006; Eenshuistra et al., 2004; Klein et al., 2000; Nieuwenhuis et al., 2000; Olincy et al., 1997; Olk and Kingstone, 2009; Raemaekers et al., 2006; Sweeney et al., 2001).

Because none of the identified studies reported effects sizes we estimated effect size for each study using reported mean reaction times and standard deviations, F-values and correlation values using the esc package (Lipsey et al., 2001). The average effect size was 1.35, CI [1.0511; 1.6527] and the effect size for our pilot study was equal to Cohen's d = 0.77. To conduct the Bayesian meta analysis, we used the R package metaBMA (Heck and Gronau, 2017). Since publication bias overinflates published estimates of effect sizes (Franco et al., 2016; Ioannidis, 2005), we based our power analysis on the lowest estimate of the effect size for the differences in reaction time between young and old group ($\delta = 0.6$). Considering that the data to be used in this study is currently being recorded in our laboratory in the context of a larger project with a fixed number of participants (see Dataset description), we used the simulation-based approach analysis design from (Schönbrodt and Wagenmakers, 2017) using the BFDA package (Schhöbrodt and Wagenmakers, 2017). In our case, assuming an effect of $\delta = 0.6$ and sample size equal to n = 156, simulation results showed that 0.5% of all simulated studies point towards the null hypothesis which specified the absence of an effect, that is, H0 of $\delta = 0$ (the rate of false negative evidence). Conversely, 92% of simulated studies show support in favor of true positive results (H1 of $\delta > 0.6$). The remaining 7.5% of simulated studies yielded inconclusive evidence. Evidence thresholds were defined at lower bound 1/6 and upper bound 6 (as proposed in the guidelines for the BFDA package (Schönbrodt and Wagenmakers, 2017).

Sample Description: Inclusion and Exclusion Criteria

Inclusion criteria for participation in the study are left and right handedness, healthy male and female participants, with an age between 20-35 years (young participants) and 60-80 (old participants). Exclusion criteria for participation are: suffering from psychiatric symptoms, severe neurological disorders (like epilepsy) or prior head injuries, a stroke, a transient circulatory disorder of the brain, diagnosis of dementia (Mini-Mental State Examination score < 26), Huntington's disease, Parkinson's disease, sensory and/or motor problems that interfere with computer tasks (e.g., the operation of a mouse), current use of psychotropic drugs (such as antidepressants, alpha-agonists, neuroleptics, mood stabilizers), intake of recreational synthetic or natural drug. Moreover, we tested the feasibility of the paradigm with the training trials and participants with bad vision either wear contact lenses or glasses.

Furthermore, data recorded from participants of the study will be excluded from the analysis in the following criteria are met: incomplete data (i.e. missing data recording from the second session), eye tracker calibration failure, i.e. more than one visual degrees deviation on average across 9 random visual stimulus presentations, less than 50% correct responses overall, more than 50% of trials rejected (see Output Measures for trial exclusion criteria). Moreover, for higher precision of the calibration and validation results, we use a small target sticker placed on the participant’s forehead, which allows head movement compensation even during blinks.

Experimental Procedure and Data Acquisition

The experiment takes place in a sound-attenuated and darkened room. The participant is seated at a distance of 68 cm from a 24-inch monitor (ASUS ROG, Swift PG248Q, display dimensions 531 × 299 mm, resolution 800 × 600 pixels resulting in a display: 400 × 298.9 mm, vertical

refresh rate of 100 Hz). Participants complete the tasks sitting alone, while research assistants monitor their progress in the adjoining room. An infrared video-based eye tracker (EyeLink 1000 Plus, SR Research, http://www.sr-research.com/) positioned next to the monitor is used to record eye position at a sampling rate of 500 Hz and an instrumental spatial resolution of 0.01°. A stable head position of the participant is ensured via a chin rest and via experimenter's instruction to stay as still as possible during data recordings. The eye tracker is calibrated and validated with a 9-point grid before each experimental block. In a validation step, the calibration is repeated until the average error for all points will be less than 1°. The eye-tracking device is recalibrated after every experimental block of the experiment (consisting of either 60 prosaccade trials or 40 antisaccade trials, see below).

In addition, EEG data is recorded (data not analyzed here). The experimental stimuli are based on an internationally standardised protocol for antisaccade testing, allowing comparisons between different labs and clinics (Antoniades et al., 2013).

The experiment was programmed in MATLAB 2016b, using the PsychToolbox extensions (Brainard, 1997; Pelli, 1997). Visual stimuli consist of horizontally arranged stimuli, targets presented on the screen are of a high contrast ratio (i.e. 11.05) in order to minimise issues related to light-adaptation level. Each trial starts with a central fixation square (visual angle of 0.6319°). Subsequently, a black square (visual angle of 0.6319°) is presented on a grey background for 1000 ms. To avoid excessive head movements (John Leigh and Zee, 2006), stimuli are always presented at the same vertical height and offset from the center (with an amplitude of 10° from the screen center). In prosaccade trials participants are instructed to perform a saccade to the peripheral stimulus - the black square presented laterally, and in antisaccade trials to perform a saccade to a corresponding location at the opposite side of the screen. The next trial starts 1000-3500 ms after the target fixations of the pro- or antisaccade. Stimuli are presented in equal numbers to the left and right side of the screen (20 per side in

the antisaccade condition and 30 per side in the control, prosaccade condition). On each experimental trial, the location (left or right) of the peripheral stimulus is randomly assigned. The standardised test protocol (Antoniades et al., 2013) consists of three blocks for the antisaccade task (40 trials per block) and two blocks of the prosaccade task (60 trials per block, control task, see Figure 1A), presented in PAAAP order to account for time-dependent effects. Before the first prosaccade block 10 practice trials, and before the first antisaccade block 5 practice trials are presented. Practice trials are aimed at acquainting the participant with our experimental procedures and were not statistically analyzed.

Each participant completes two recording sessions in a test-retest experimental design with an interval of one week (acceptable range: 7-9 days) between recording sessions (at the same time of day). During both visits, the same experimental protocol is followed, including the same order of tasks.

Eye-Tracking Data Preprocessing

The EyeLink 1000 tracker computes eye-position data, measures pupil diameter and identifies events such as saccades, fixations, and blinks. Saccade onsets are detected using the eye tracking software default settings: acceleration larger than 8000° per sec², a velocity above 30° per sec, and a deflection above 0.1°. We extract the following information about the saccades: start and end time, duration, coordinates of start positions and end positions on the computer screen in pixels, amplitudes, and eye velocity.

Fixations are defined as time periods without saccades and eye blinks are regarded as a special case of a fixation, where the pupil diameter is either zero or outside a dynamically computed valid pupil. Thus fixation might include small saccades (i.e. microsaccades), which fall below the threshold for saccade detection. In the present study, we focus only on standard

saccades (not microsaccades), because all considered output measures are based on these standard saccades."

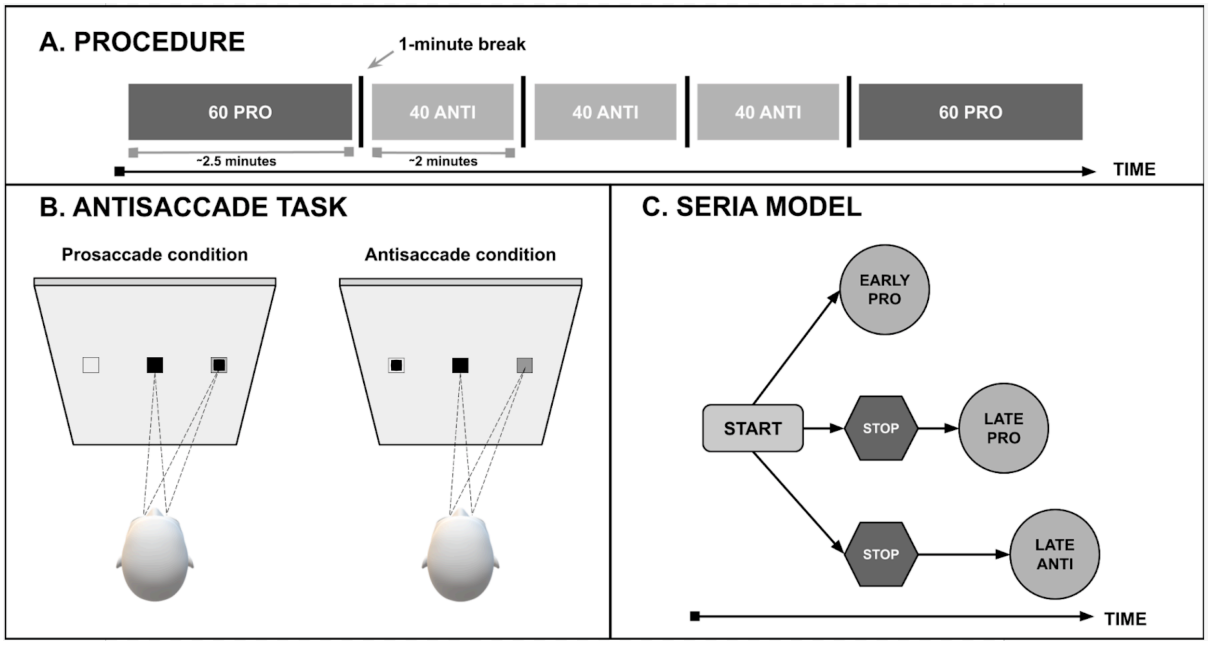

Figure 1: (A) the experimental procedure of a single run, consisting of prosaccade task (PRO) and antisaccade task (ANTI) blocks, which each consisted of either 40 or 60 trials per block. There was a 1-min between each block. (B) Schematic top view of the experimental setup and gaze behavior during a prosaccade and antisaccade condition trial. The black square represents the target fixation, and the black dot represents the peripheral stimulus. The peripheral stimulus is presented 1000 ms on the screen and starts after a duration of the target fixation of 800-1200 ms. (C) The sequence of latent events assumed by the Stochastic Early Reaction, Inhibition and Late Action (SERIA) Model, generating as output either early prosaccades (EARLY PRO), late prosaccades (LATE PRO), or antisaccade events (LATE ANTI).

Output Measures

The output measures of interests are: *Reaction time for the first saccade*, defined as time from onset of the peripheral stimulus to the start of the saccade (Antoniades et al., 2013), irrespective of whether the saccade was elicited in the correct direction. An *error* is defined as a saccade towards the stimulus in an antisaccade block, and away from the stimulus in a prosaccade

block. The *error rate* for each participant will be calculated as the proportion of erroneous trials to all valid trials separately for anti- and prosaccade blocks. Additionally, we will extract *peak saccadic velocity* for each saccade as provided by eye tracker recordings. The *gain of the first saccade* will be calculated as a ratio of actual saccade amplitude divided by the desired saccade amplitude (in our experimental setup equal to 10 deg, based on Antoniades et al., 2013). Trial exclusion criteria are based on Antoniades et al. (2013): occurrences of eye blinks between the cue presentation and the saccade, reaction times of less than 50 ms duration, a saccade onset later than 800 ms after cue presentation. If 50% or more trials were rejected the subject was excluded.

Data analysis

The two primary goals of our study are testing the presence of age differences in all outcome measures and inspecting their reliability across the two test-retest recording sessions. For each of the goals, we describe below the planned analysis pipeline, including all preprocessing steps and planned analyses.

Age Differences

The presence of age differences in all outcome measures (reaction times, error rates, peak saccadic velocities, saccade gains, two model parameters of SERIA: inhibitory fail probability and inhibitory fail reaction time (see section "Computational model" for description of model parameters) will be investigated. Single trials that were not excluded during preprocessing (see "Output measures" for trial exclusion criteria) from all subjects will be used for fitting a multivariate Bayesian generalized linear mixed model. We will use the brms package which offers robust estimates in the context of multilevel modelling (Bürkner, 2018, 2017; Kozak and Cuthbert, 2016). To improve convergence and guard against overfitting, we will use weakly

informative Cauchy priors in line with the recommendations for Bayesian regression models (Gelman et al., 2008). We will use the data from both time points and random intercepts will be added for the Participant factor. The predictor Type (levels: antisaccade condition, prosaccade condition) will be included to account for the influence of the Type of the experimental block. The model will fit at the same time the four dependent variables (reaction times, error rates, peak saccadic velocities, saccade gains). To account for possible multiple comparisons, we will correct the effective number of tests using the approach of Nyholt (2004), which, based on the ratio of observed eigenvalue variance to its maximum, gives the proportional reduction in the number of variables in a set, and therefore provides a useful alternative to more computationally intensive permutation tests. Then, we will report the adjusted alpha level of the Bayesian posterior credibility intervals (CI).

Test- Retest Reliability

In order to quantify test-retest reliability for the output measures collected at the two recording sessions per subject, we will calculate one-way random effects model intraclass correlation coefficients (ICCs) using the absolute agreement measure among multiple observations (Bhapkar, 1966; Finn, 1970; McGraw and Wong, 1996), with the open source software package irr (https://CRAN.R-project.org/package=irr) for reaction times, peak saccade velocity, error rates and gain of the first saccade and the quantities obtained from the computational model. We will use the following, generally adopted interpretation of ICC, introduced by (Cicchetti, 1994): Less than 0.40 (poor reliability), between 0.40 and 0.59 (fair reliability), between 0.60 and 0.74 (good reliability) and between 0.75 and 1.00 (excellent reliability).

Additionally, we will also use Bland-Altman plots (Bland and Altman, 1999) for graphical comparison of two measurements from test and retest recording sessions. In the Bland-Altman plot each sample is represented on the graph by plotting the mean value of the two assessments

against the difference value between them. The chart can then highlight possible anomalies, such as revealing that one method overestimates high values and underestimates low values (Kalra, 2017). We will also use a quantitative method assessing the agreement of test and retest (first and second measurement). It's based on a priori defined limits of agreement (as for other relevant measures, it was recommended that 95% of the data points should lie within ±1.96 SD of the mean difference – limits of agreement; Earthman, 2015; Sedgwick, 2013).

Computational model

We will use the PRO- Stop-Antisaccade (PROSA) and the Stochastic Early Reaction, Inhibition, and late Action model (SERIA) model (Aponte et al., 2017) to fit experimental data from the antisaccade task to estimate latent, not directly observable processes. PROSA and SERIA are inspired by the hypothesis that antisaccades are the result of competing decision mechanisms that interact nonlinearly with each other. This approach is based on previous proposals and fits the to-be explained reaction time and error rate in the double step and search step tasks (Noorani and Carpenter, 2013). SERIA and PROSA offer a formal, probabilistic approach to the antisaccade task and provide detailed information about the participants' performance.

Briefly, the PROSA model assumes that the reaction time and the response (either pro or antisaccade) in a given trial are caused by the interaction of three competing processes: eliciting a prosaccade, inhibitory command to stop a prosaccade, and eliciting an antisaccade. On the other hand, in the SERIA model, four different units can be distinguished: the early prosaccade unit, the inhibitory unit (that can stop early prosaccades), the antisaccade unit, and the late prosaccade unit (see Figure 1C for an illustration of the model). The exact details of The PROSA and SERIA are described in Aponte et al. (2017).

We will use the SEM toolbox (Aponte et al., 2017) and the method for model fitting used by Aponte et al. (2017), based on the Metropolis-Hastings algorithm (Gelman et al., 2003). Moreover, we will apply a hierarchical method of fitting the model, which treats the group mean as prior to the parameters and therefore offers a form of regularization based on observations from the population. Our data (only valid trials, see: Sample Description: Inclusion and Exclusion Criteria) will be entered into the models as a structure with fields representing the reaction time and the corresponding action (either pro- or antisaccade). The result is an array of samples from the target distribution, which can be used to compute summary statistics. To compare the PROSA and SERIA model we will extract the model-based fits and compare obtained histograms, computing the means and standard deviations of the reaction times and the fits, between old and young participants' groups.

Subsequently, we will compare PROSA and SERIA fits for young and the old participants, based on obtained model evidence, as described in (Aponte et al., 2017).

Pilot Data

The primary purpose of the pilot data analysis was to assure that our test-retest experimental design is a stable and reliable method to further testing age differences. According to our power analysis (see section Methods), the pilot data set is underpowered, and thus, we did not conduct any statistical tests on it. Instead, we present the raw distributions and reciprobit plots of reaction times. Additionally, we include ICCs for four output measures and Bland-Altman plots for reaction times and error rates, which need to be interpreted with caution, because of the small sample size (methods for obtaining them are described in the methods section).

Participants

Data for the pilot study were recorded from 22 healthy young subjects (20-25 years, mean age 23.6 years, sd = 3.3 years) and 22 healthy elderly subjects (>60 years, mean age 68.9 years, sd = 2.9 years). Data from four participants were discarded due to low performance in the antisaccade task (error rate > 50%). The final sample used for pilot data analysis thus consists of 40 participants.

Results

Output measures

Across all 40 subjects, a total of 19'200 trials were recorded, from which 906 trials were excluded based on the trial exclusion criteria described in the Methods section. Out of the total 906 excluded trials, 288 were occurrences of eye blinks between the cue presentation and the saccade, 526 had reaction times of less than 50 ms duration, and 92 had a saccade onset later than 800 ms after cue presentation. For each experimental trial we extracted: reaction time for the first saccade, information if the participant looked in the correct direction or not, peak saccadic velocity, gain of the first saccade. Table 4 compares the results obtained from the pilot data set. Descriptives of each of the extracted measures are presented separately for pro- and antisaccades, young and old participants.

	Young Group (n=20)								Old Group (n=20)							
	Prossacade Condition				Antisaccade Condition				Prossacade Condition				Antisaccade Condition			
	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max	mean	sd	min	max
Reaction time (ms)	268	83	51	790	303	88	51	786	309	118	51	796	360	130	51	794
Error rate (%)	1.3	1.92	0	10	7.83	6.54	0	27.5	5.35	5.31	0	21.6	17.16	14.7	0	57.4
Gain of the saccade	0.81	0.18	0.01	2.58	0.79	0.22	0.01	3.48	0.76	0.28	0.01	3.07	0.7	0.32	0.01	4.03
Peak saccadic velocity (deg/s)	331	229	45	3270	326	259	5.0	3270	288	193	44.0	3270	267	210	44	3270

Table 1: Descriptives of reaction times for the first saccade, error rate, gain of the first saccade (ratio of actual saccade amplitude divided by the desired saccade amplitude), and peak saccadic velocity for the pro- and antisaccade condition for the young and old group.

To assess the contribution of different factors to an experiment's results (Carpenter et al., 2007; Noorani and Carpenter, 2013), we used reciprobit plots, as recommended in the internationally standardized antisaccade protocol (Antoniades et al., 2013). Figure 2 shows data distributions of all trials from the young group (left part of the figure) and the old group (right part of the figure). In the antisaccade task, the latency distributions of correct antisaccades and error prosaccades have characteristics that are different from those seen in the control (prosaccade)

condition. The error responses were slightly delayed for the antisaccade as compared to the prosaccade condition (especially evident in the old participants), and it is visible that there were far fewer errors for prosaccades than for antisaccades.

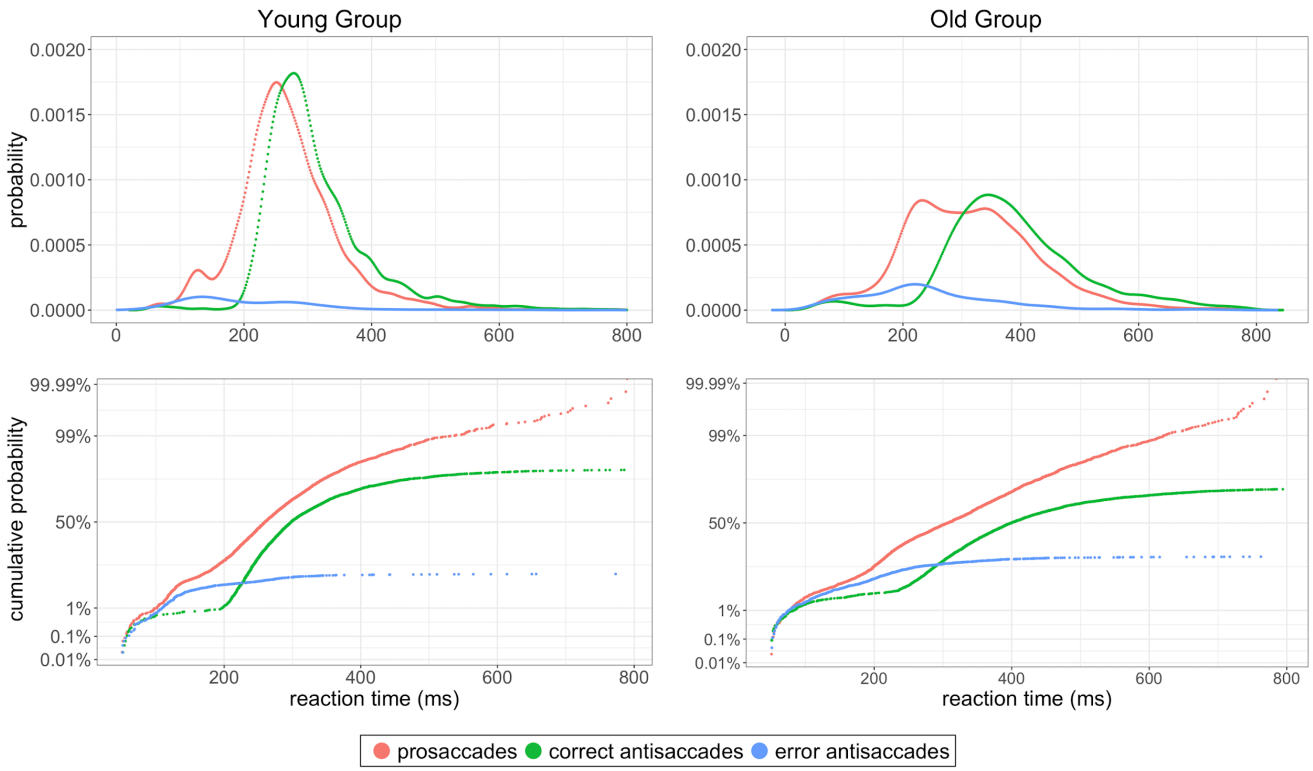

Figure 2: Top panels: Raw distributions with error responses plotted as a cumulative proportion of the total number of trials for young and old group, showing a rightward shift of the correct antisaccade distribution relative to both the prosaccade and error antisaccades distributions. Bottom panels: The same data as shown above as reciprobit plots. Error responses are plotted as a cumulative proportion of the total number of trials.

Test- Retest Reliability

Our pilot study confirmed the high test-retest reliability for reaction times, first saccade gains and peak saccadic velocity (see Table 2). A possible explanation for the low ICCs for error rates of young participants might be that error rates, especially for the prosaccade task are low (<5% of all trials), and thus, we had not enough data to obtain stable estimates for this output measure. Figure displays the reciprobit plot of the error rate, separately for old and young group of participants, for test and retest.

Table 2. Intraclass correlation coefficients with 95%-confidence intervals in brackets for four output measures, separately for pro- and antisaccade condition and for old and young group

	Young Group (n=20)		Old Group (n=20)	
	Prosaccades	Antisaccades	Prosaccades	Antisaccades
Reaction time	0.66 (0.51; 0.77)	0.64 (0.53; 0.71)	0.85 (0.78; 0.9)	0.8 (0.74; 0.85)
Error rate	0.22 (0.09; 0.41)	0.45 (0.33; 0.56)	0.47 (0.27; 0.62)	0.75 (0.67; 0.8)
Gain of the saccade	0.51 (0.32; 0.65)	0.62 (0.52; 0.7)	0.64 (0.49; 0.75)	0.61 (0.5; 0.7)
Peak saccadic velocity	0.51 (0.33; 0.66)	0.5 (0.39; 0.61)	0.71 (0.58; 0.8)	0.59 (0.48; 0.69)

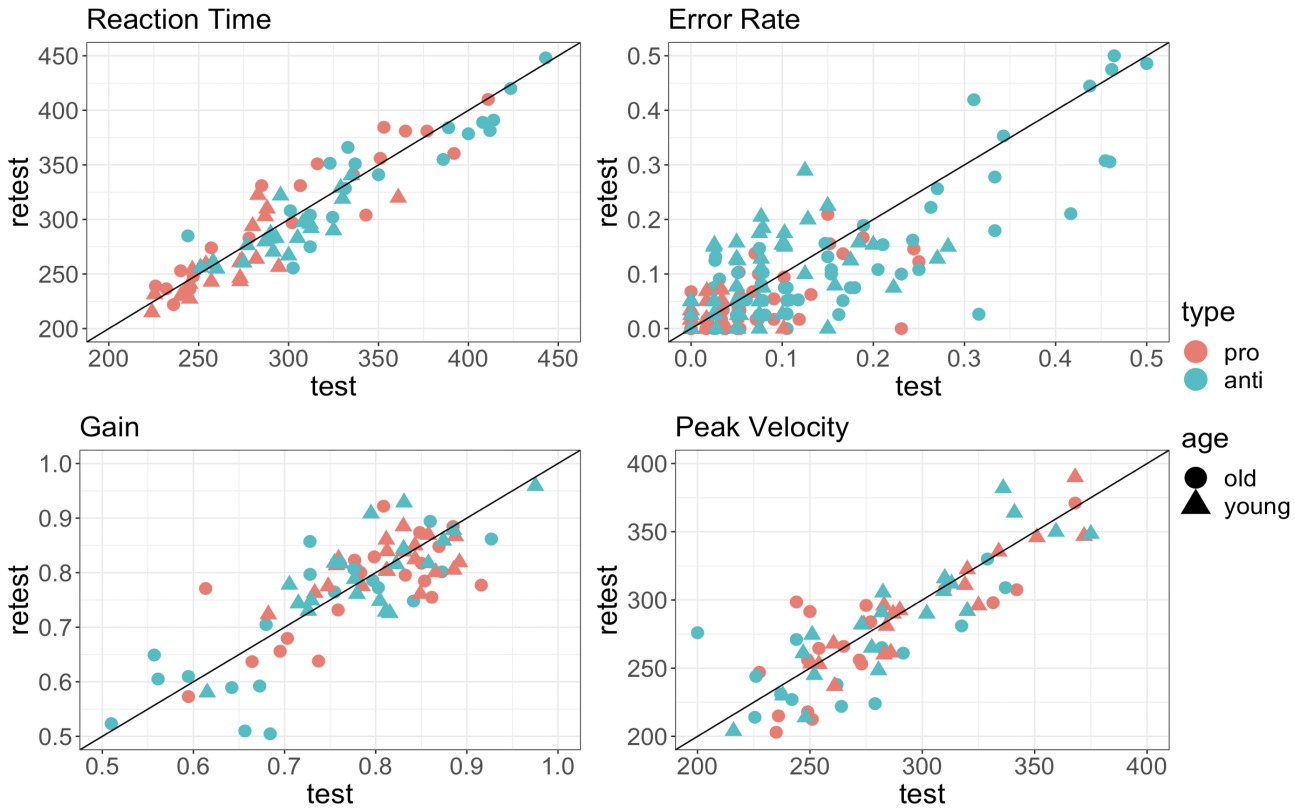

Figure 3: Paired distributions of four output measures (reaction time, error rate, gain, peak velocity) for test and retest measurement timepoints. Each point represents one subject. The solid black line corresponds to an ideal correlation with a fixed slope of 1.

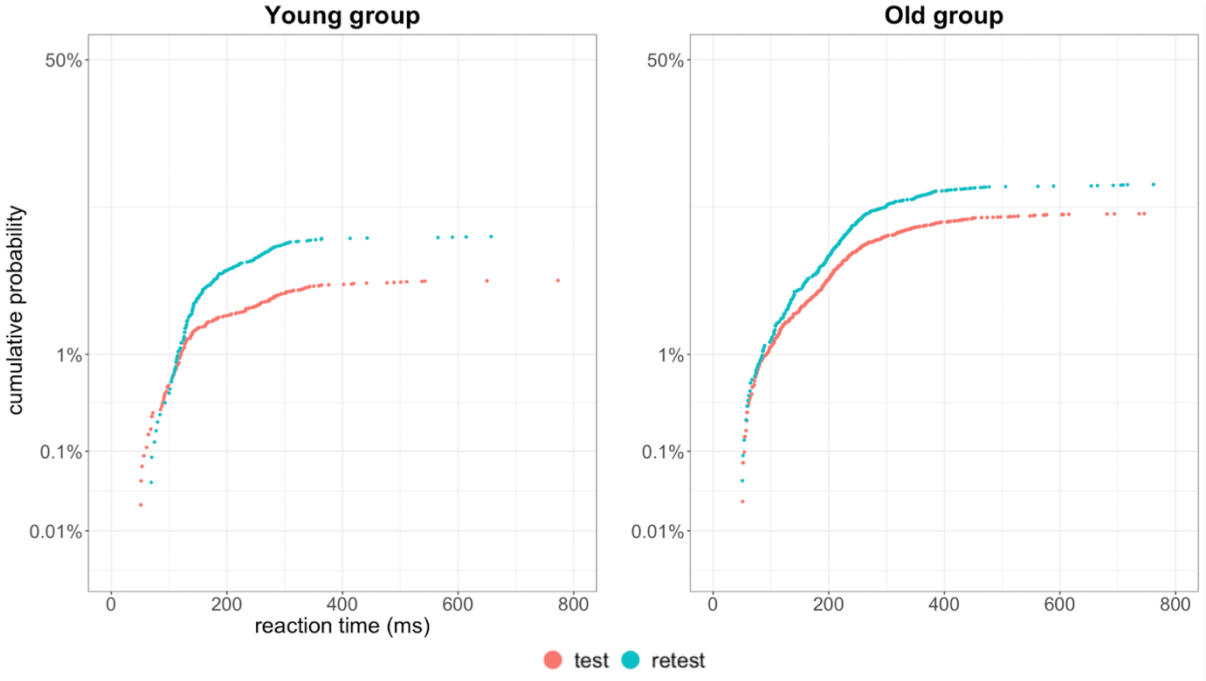

Figure 4: Reciprobit plots for error rate in the antisaccade trials, comparison for the young and old group, for test and retest. Error responses are plotted as a cumulative proportion of the total number of trials.

Additionally, Bland-Altman plots were used to graphically represent the agreement between the two measurements. According to Kalra (2017), 95% of the data points should lie within ±1.96 SD of the mean difference limits of agreement. From the data in Figure 4, it is apparent that our study design can provide reliable results and is suitable for further testing in the main study, with a larger sample size.

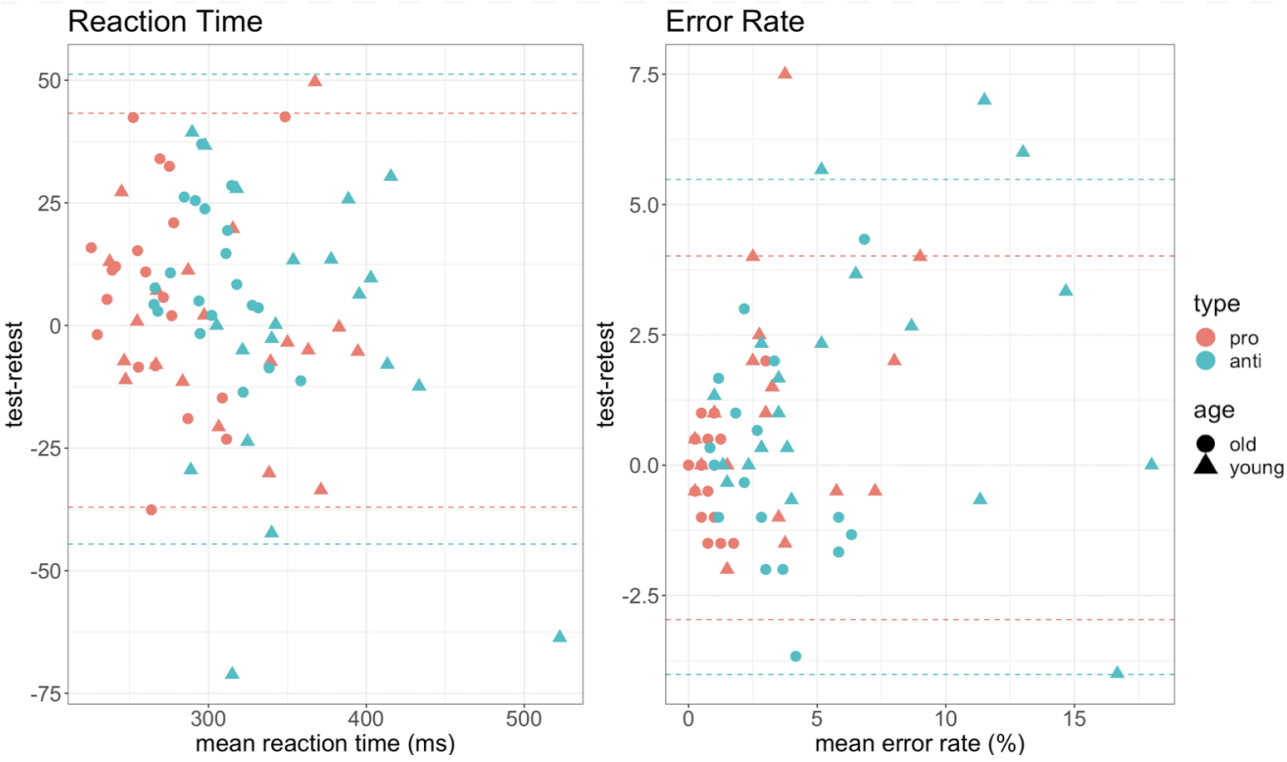

Figure 5: Bland-Altman plots for two measures of interest: error rate and reaction time. Horizontal lines are drawn at the limits of agreement, which are defined as the mean difference plus and minus 1.96 times the standard deviation of the differences.

References

Abel LA, Douglas J (2007) Effects of age on latency and error generation in internally mediated saccades. Neurobiol Aging 28:627–637.

Antoniades C, Ettinger U, Gaymard B, Gilchrist I, Kristjánsson A, Kennard C, John Leigh R, Noorani I, Pouget P, Smyrnis N, Tarnowski A, Zee DS, Carpenter RHS (2013) An internationally standardised antisaccade protocol. Vision Res 84:1–5.

Aponte EA, Schöbi D, Stephan KE, Heinzle J (2017) The Stochastic Early Reaction, Inhibition, and late Action (SERIA) model for antisaccades. PLoS Comput Biol 13:e1005692.

Aponte EA, Stephan KE, Heinzle J (2019) Switch costs in inhibitory control and voluntary behaviour: A computational study of the antisaccade task. European Journal of Neuroscience.

Bhapkar VP (1966) A Note on the Equivalence of Two Test Criteria for Hypotheses in Categorical Data. Journal of the American Statistical Association.

Bialystok E, Craik FIM, Ryan J (2006) Executive control in a modified antisaccade task: Effects of aging and bilingualism. Journal of Experimental Psychology: Learning, Memory, and Cognition.

Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. Stat Methods Med Res 8:135–160.

Bojko A, Kramer AF, Peterson MS (2004) Age equivalence in switch costs for prosaccade and antisaccade tasks. Psychol Aging 19:226–234.

Brainard DH (1997) The Psychophysics Toolbox. Spatial Vision.

Bürkner P-C (2018) Advanced Bayesian Multilevel Modeling with the R Package brms. The R Journal.

Bürkner P-C (2017) brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software.

Butler KM, Zacks RT (2006) Age deficits in the control of prepotent responses: evidence for an inhibitory decline. Psychol Aging 21:638–643.

Butler KM, Zacks RT, Henderson JM (1999) Suppression of reflexive saccades in younger and older adults: age comparisons on an antisaccade task. Mem Cognit 27:584–591.

Camalier CR, Gotler A, Murthy A, Thompson KG, Logan GD, Palmeri TJ, Schall JD (2007) Dynamics of saccade target selection: race model analysis of double step and search step saccade production in human and macaque. Vision Res 47:2187–2211.

Carpenter RHS, R H, McDonald SA (2007) LATER predicts saccade latency distributions in reading. Experimental Brain Research.

Cicchetti DV (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 6:284–290.

Connelly SL, Lisa Connelly S, Hasher L (1991) Aging and the inhibition of spatial location. PsycEXTRA Dataset.

Crawford TJ, Higham S, Renvoize T, Patel J, Dale M, Suriya A, Tetley S (2005) Inhibitory control of saccadic eye movements and cognitive impairment in Alzheimer's disease. Biol Psychiatry 57:1052–1060.

Crawford TJ, Smith ES, Berry DM (2017) Eye Gaze and Aging: Selective and Combined Effects of Working Memory and Inhibitory Control. Front Hum Neurosci 11:563.

Cutsuridis V, Smyrnis N, Evdokimidis I, Perantonis S (2007) A neural model of decision-making by the superior colicullus in an antisaccade task. Neural Netw 20:690–704.

Earthman CP (2015) Body Composition Tools for Assessment of Adult Malnutrition at the Bedside: A Tutorial on Research Considerations and Clinical Applications. JPEN J Parenter Enteral Nutr 39:787–822.

Eenshuistra RM, Richard Ridderinkhof K, van der Molen MW (2004) Age-related changes in antisaccade task performance: Inhibitory control or working-memory engagement? Brain and Cognition.

Feng G (2012) Is there a common control mechanism for anti-saccades and reading eye

movements? Evidence from distributional analyses. Vision Research.

Finn RH (1970) A Note on Estimating the Reliability of Categorical Data. Educational and Psychological Measurement.

Franco A, Malhotra N, Simonovits G (2016) Underreporting in Psychology Experiments. Social Psychological and Personality Science.

Gelman A, Carlin JB, Stern HS, Rubin DB (2003) Hierarchical models. Bayesian data analysis 120–160.

Gelman A, Jakulin A, Pittau MG, Su Y-S (2008) A weakly informative default prior distribution for logistic and other regression models. Ann Appl Stat 2:1360–1383.

Hallett PE (1978) Primary and secondary saccades to goals defined by instructions. Vision Res 18:1279–1296.

Houx PJ, Vreeling FW, Jolles J (1993) Decline of Psychomotor Performance: Calender or Health? Sensorimotor Impairment in the Elderly.

Ioannidis JPA (2005) Why most published research findings are false. PLoS Med 2:e124.

John Leigh R, Zee DS (2006) The Neurology of Eye Movements. Oxford University Press, USA.

Kalra A (2017) Decoding the Bland–Altman plot: Basic review. Journal of the Practice of Cardiovascular Sciences.

Kanasi E, Ayilavarapu S, Jones J (2016) The aging population: demographics and the biology of aging. Periodontol 2000 72:13–18.

Klein C, Fischer B, Hartnegg K, Heiss WH, Roth M (2000) Optomotor and neuropsychological performance in old age. Exp Brain Res 135:141–154.

Kozak MJ, Cuthbert BN (2016) The NIMH Research Domain Criteria Initiative: Background, Issues, and Pragmatics. Psychophysiology 53:286–297.

Lipsey MW, Wilson DB, David Wilson, MS RN C(inc) (2001) Practical Meta-Analysis.

Chronicle Books.

Logan GD, Cowan WB, Davis KA (1984) On the ability to inhibit simple and choice reaction time responses: A model and a method. Journal of Experimental Psychology: Human Perception and Performance.

McGraw KO, Wong SP (1996) Forming inferences about some intraclass correlation coefficients. Psychol Methods 1:30.

Meeter M, Van der Stigchel S, Theeuwes J (2010) A competitive integration model of exogenous and endogenous eye movements. Biol Cybern 102:271–291.

Nieuwenhuis S, Ridderinkhof KR, de Jong R, Kok A, van der Molen MW (2000) Inhibitory inefficiency and failures of intention activation: age-related decline in the control of saccadic eye movements. Psychol Aging 15:635–647.

Noorani I, Carpenter RHS (2013) Antisaccades as decisions: LATER model predicts latency distributions and error responses. Eur J Neurosci 37:330–338.

Nyholt DR (2004) A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. Am J Hum Genet 74:765–769.

Olincy A, Ross RG, Youngd DA, Freedman R (1997) Age diminishes performance on an antisaccade eye movement task. Neurobiol Aging 18:483–489.

Olk B, Kingstone A (2009) A new look at aging and performance in the antisaccade task: The impact of response selection. European Journal of Cognitive Psychology.

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spatial Vision.

Peltsch A, Hemraj A, Garcia A, Munoz DP (2011) Age-related trends in saccade characteristics among the elderly. Neurobiol Aging 32:669–679.

Raemaekers M, Vink M, van den Heuvel MP, Kahn RS, Ramsey NF (2006) Effects of aging

on BOLD fMRI during prosaccades and antisaccades. J Cogn Neurosci 18:594–603.

Ramat S, Leigh RJ, Zee DS, Optican LM (2007) What clinical disorders tell us about the neural control of saccadic eye movements. Brain 130:10–35.

Reuter B, Rakusan L, Kathmanna N (2005) Poor antisaccade performance in schizophrenia: an inhibition deficit? Psychiatry Res 135:1–10.

Rey-Mermet A, Gade M (2018) Inhibition in aging: What is preserved? What declines? A meta-analysis. Psychon Bull Rev 25:1695–1716.

Rey-Mermet A, Meier B (2017) How long-lasting is the post-conflict slowing after incongruent trials? Evidence from the Stroop, Simon, and flanker tasks. Atten Percept Psychophys 79:1945–1967.

Schhnbrodt FD, Wagenmakers E-J (n.d.) Bayes Factor Design Analysis: Planning for Compelling Evidence. SSRN Electronic Journal.

Schönbrodt FD, Wagenmakers E-J (n.d.) Bayes Factor Design Analysis: Planning for compelling evidence.

Sedgwick P (2013) Limits of agreement (Bland-Altman method). BMJ.

Shafiq-Antonacci R, Maruff P, Masters C, Currie J (2003) Spectrum of saccade system function in Alzheimer disease. Arch Neurol 60:1272–1278.

Spieler DH, Balota DA, Faust ME (1996) Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. J Exp Psychol Hum Percept Perform 22:461–479.

Sweeney JA, Rosano C, Berman RA, Luna B (2001) Inhibitory control of attention declines more than working memory during normal aging. Neurobiol Aging 22:39–47.

Verhaeghen P, Cerella J (2002) Aging, executive control, and attention: a review of meta-analyses. Neurosci Biobehav Rev 26:849–857.

Wiecki TV, Frank MJ (2013) A computational model of inhibitory control in frontal cortex

and basal ganglia. Psychol Rev 120:329–355.