University of Groningen

Stimulus contrast, pursuit mode, and age strongly influence tracking performance on a continuous visual tracking task

Vrijling, Anne; de Boer, Minke; Renken, Remco; Marsman, Jan-Bernard C; Grillini, Alessandro; Petrillo, Enrico; Heutink, Joost; Jansonius, Nomdo M.; Cornelissen, Frans

Published in:
Vision Research

DOI:
10.1016/j.visres.2023.108188

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2023

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Vrijling, A., de Boer, M., Renken, R., Marsman, J.-B. C., Grillini, A., Petrillo, E., Heutink, J., Jansonius, N. M., & Cornelissen, F. (2023). Stimulus contrast, pursuit mode, and age strongly influence tracking performance on a continuous visual tracking task. *Vision Research*, *205*, Article 108188. https://doi.org/10.1016/j.visres.2023.108188

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 21-04-2025

Vision Research 205 (2023) 108188

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Stimulus contrast, pursuit mode, and age strongly influence tracking performance on a continuous visual tracking task

A.C.L. Vrijling [a,b,1,*], M.J. de Boer [a,1], R.J. Renken [c], J.B.C. Marsman [c], A. Grillini [d], C.E. Petrillo [d], J. Heutink [b,e], N.M. Jansonius [a,f], F.W. Cornelissen [a]

[a] *Laboratory of Experimental Ophthalmology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands*
[b] *Royal Dutch Visio, Centre of Expertise for Blind and Partially Sighted People, Huizen, the Netherlands*
[c] *Cognitive Neuroscience Center, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands*
[d] *Reperio BV, Groningen, the Netherlands*
[e] *Department of Clinical and Developmental Neuropsychology, University of Groningen, Groningen, the Netherlands*
[f] *Department of Ophthalmology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands*

ARTICLE INFO

Keywords:
Eye movements
Continuous stimulus tracking
Visual field
Psychophysics
Weber contrast

ABSTRACT

Human observers tend to naturally track moving stimuli. This tendency may be exploited towards an intuitive means of screening visual function as an impairment induced reduction in stimulus visibility will decrease tracking performance. Yet, to be able to detect subtle impairments, stimulus contrast is critical. If too high, the decrease in performance may remain undetected. Therefore, for this approach to become reliable and sensitive, we need a detailed understanding of how age, stimulus contrast, and the type of stimulus movement affect continuous tracking performance. To do so, we evaluated how well twenty younger and twenty older participants tracked a semi-randomly moving stimulus (Goldmann size III, 0.43 degrees of visual angle), presented at five contrast levels (5%-10%-20%-40%-80%). The stimulus could move smoothly only (smooth pursuit mode) or in alternation with displacements (saccadic pursuit mode). Additionally, we assessed static foveal and peripheral contrast thresholds. For all participants, tracking performance improved with increasing contrast in both pursuit modes. To reach threshold performance levels, older participants required about twice as much contrast (20% vs. 10% and 40% vs. 20% in smooth and saccadic modes respectively). Saccadic pursuit detection thresholds correlated significantly with static peripheral contrast thresholds (rho = 0.64). Smooth pursuit detection thresholds were uncorrelated with static foveal contrast thresholds (rho = 0.29). We conclude that continuous visual stimulus tracking is strongly affected by stimulus contrast, pursuit mode, and age. This provides essential insights that can be applied towards new and intuitive approaches of screening visual function.

1. Introduction

Accurate assessment of visual function is essential in ophthalmic care and rehabilitation. One of the aspects to assess is the integrity of the visual field. The current gold standard for this is standard automated perimetry (SAP). During SAP, stimuli at different luminances are systematically presented at various test locations throughout the visual field. Meanwhile, the participant has to maintain fixation on a central target and decide on the presence of a stimulus. If so, their conscious decision regarding the presence of a stimulus has to be translated into a manual action (pressing a button). Therefore, performing SAP requires the understanding of relatively long and complicated verbal test instructions, multi-tasking, and prolonged focused attention (for over 6 min per eye) (Bengtsson et al., 1998). This renders data quality dependent on participant compliance and abilities. These requirements can complicate the use of SAP in elderly individuals, children, and individuals with cognitive and/or motor impairments. This limits the availability of reliable information on their visual functioning which, in turn, may limit the quality of ophthalmic and rehabilitation care. In such participant groups, easier and more intuitive ways to screen visual function could potentially contribute to improving the quality of care for these individuals.

* Corresponding author at: University Medical Center Groningen, Department of Ophthalmology, PO BOX 30.001, 9700 RB Groningen, the Netherlands.
E-mail address: a.c.l.vrijling@umcg.nl (A.C.L. Vrijling).
[1] shared first authorship

https://doi.org/10.1016/j.visres.2023.108188
Received 24 May 2022; Received in revised form 23 January 2023; Accepted 23 January 2023
Available online 9 February 2023
0042-6989/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Previous studies have shown that eye movements can serve as an alternative to manual responses in visual function assessment (Grillini et al., 2018, 2020, 2021; Jones et al., 2019; Kadavath Meethal et al., 2018; Mazumdar et al., 2014, 2019, 2020; McTrusty et al., 2017; Murray et al., 2009; Pel et al., 2013). Moreover, studies measuring saccadic response times to peripheral targets have shown that such eye movement-based responses can be used to detect glaucomatous visual field defects (Kadavath Meethal et al., 2018; Mazumdar et al., 2014; McTrusty et al., 2017; Pel et al., 2013). When used in this way, eye tracking simplifies the approach, yet it still requires the participant to fixate on a target and to keep fixating until the next target is perceived. Such fixating is a rather uncommon requirement in natural vision. Instead, human observers have the natural tendency to move their eyes and track moving or appearing stimuli. Recently, we took an approach in which we try to exploit these natural tendencies for visual function screening. In this approach, participants are asked to track a continuously moving stimulus with their eyes (Grillini et al., 2018). This approach is potentially suitable as a screening tool as a reduced stimulus visibility will result in a momentarily decrease in tracking performance. Such a reduced visibility could be due to a reduced contrast sensitivity caused by a visual field defect or neuro-ophthalmic deficit (Soans et al., 2021). Even though a continuous tracking task requires stabilizing visual information on one's retina, it is very intuitive.

In our continuous tracking task, contrast sensitivity is only indirectly assessed via its effect on tracking behavior. Good tracking requires a suprathreshold stimulus, but if the stimulus contrast is too high, performance will not or only marginally decrease with small and local reductions in contrast sensitivity. Therefore, a full understanding of the relationship between stimulus contrast and tracking performance on our specific task is required for it to become suitable as a visual function screening tool. Additionally, both contrast sensitivity (Haegerstrom-Portnoy et al., 1999; Heijl et al., 1987; Katz & Sommer, 1986; Vonthein et al., 2007) and smooth pursuit performance (Sharpe & Sylvester, 1978) decrease with age. Yet, it is unknown whether such previous findings also hold for an unpredictably moving and jumping stimulus.

Therefore, the aim of our study was to determine how continuous tracking performance in healthy participants is influenced by stimulus contrast, participant age and the type of stimulus movement. For this purpose, we evaluated the agreement between eye and stimulus movements in a group of younger and older healthy participants who had to continuously track a small, semi-randomly moving, stimulus that moved in one of two modes: 1) *smooth pursuit mode*, where the stimulus moves continuously with the direction and velocity of the motion varying in a pseudo-random manner. The participant may track this stimulus using smooth pursuit eye movements, but catch-up saccades are also allowed, and 2) *saccadic pursuit mode*, where the stimulus also moves continuously as in the smooth pursuit mode, but at pseudo-random intervals the stimulus additionally makes jumps to new test locations. Consequently, to keep track of the stimulus, the participant will need to make saccadic eye movements. In different sets of trajectories, stimuli were presented at five contrast levels covering the range from near-threshold foveally to suprathreshold peripherally (5%-80% Weber contrast, in 0.3 log steps). Additionally, we compared tracking performance to more conventionally assessed static foveal and peripheral Weber contrast thresholds.

2. Methods

2.1. Study population

We recruited 22 younger and 23 older participants. Due to technical issues (mainly calibration issues and unstable pupil detection), we could either not complete the measurements or had incomplete data in 2 younger and 3 older participants. Therefore, a total of twenty younger (17 females, age range 18–29, mean age of 21 years) and twenty older (11 females, age range 51–77, mean age of 65 years) participants with normal vision completed the study. The study was approved by the Medical Ethical Committee of the University Medical Center Groningen (NL70382.042.20). The study followed the tenets of the Declaration of Helsinki.

The dominant eye (Miles test, Roth et al. 2002) of the participant was tested. The non-tested eye was patched. The tested eye had to have a best-corrected visual acuity (BCVA) of 0.10 logMAR or better, verified with the ETDRS chart of the VIP-software (Vision Inspector Pro, Optical Diagnostics, version 3.0). Exclusion criteria were any known eye abnormalities, a positive family history of glaucoma, or neurological disorders that could affect test performance (assessed with a questionnaire). We did not perform SAP in our study population.

2.2. Experimental set-up

Stimuli were presented on a 24.5-inch IPS monitor (OptixMag25R1X, Micro-Star International Co., Ltd., New Taipei City, Taiwan), with a resolution of 1920 × 1080 pixels (49 × 29 degrees of visual angle at the applied viewing distance of 60 cm) and a framerate of 240 Hz. The luminance output of the monitor was calibrated using an automatic i1Display device (X-Rite, China). The calibration was performed on a uniform background set to RGB (gray) values 0–255, in steps of 1 Gy value. At each RGB value, the luminance output at the center of the screen was recorded. To correct for luminance differences across the screen, we measured the luminance of uniform backgrounds at gray values 0–250 in steps of 25 with a handheld luminance meter (LS-110, Konica Minolta, Inc.). At each step, the luminance output was recorded at five locations on the screen (center and all four corners) and averaged to obtain the luminance output. These 11 luminance outputs consistently differed −6% from the corresponding luminance output as measured with the i1Display device. The final luminance output was the output of the i1Display device adjusted with the −6% difference. A look-up table for conversion between cd/m^2 and RGB values was created. The minimum output of the screen was 0.4 cd/m^2 (RGB value 0) and the maximum output was 358.1 cd/m^2 (RGB value 255).

The experiment was performed in a dimly lit room with the experimental monitor being the main source of illumination. Participants were seated in front of the screen at a viewing distance of 60 cm with their head placed on a chinrest and stabilized with a forehead rest to minimize head movements. All tests were performed monocularly with optimal correction (spherical equivalent) for the viewing distance. Stimulus display and gaze recording were controlled by the Psychtoolbox and Eyelink Toolbox extensions (Brainard, 1997; Cornelissen et al., 2002; Kleiner et al., 2007; Pelli, 1997) for MATLAB (The Mathworks, Inc., version 2021a). The host PC for stimulus presentation was an HP Elitedesk 800 G3 with a NVIDIA Geforce GTX 1080 graphics card, running Windows 10.

A desktop-mounted Eyelink 1000 Plus eye-tracker (SR Research Ltd., Ottawa, Ontario, Canada), running software version 4.51, was used to measure participants' eye movements. Due to technical issues, we were forced to switch eye-tracker after 16 (14 younger, 2 older) participants. For the remaining participants, we used a desktop-mounted Eyelink Portable Duo eye-tracker (SR Research Ltd.). Both eye-trackers recorded monocular gaze data at a sampling frequency of 1000 Hz and were functionally identical for the purpose of this study. The eye-tracker model did not influence the results. At the start of the experiment, the eye-tracker was calibrated using the built-in 9-point calibration routine. Calibration accuracy was verified with the built-in validation routine. Calibration and validation were repeated until the calibration accuracy was deemed 'good' by the tracker software (average error$<0.5^\circ$ and maximum error $< 1.0^\circ$). In case such a 'good' calibration could not be achieved, the participant was excluded. If one of the nine points could not be recorded accurately (i.e., error $> 1.0^\circ$; usually for one of the upper corners), while the remaining eight points only had a small error, the experiment was continued. A drift correction was performed prior to presenting each set of trajectories. If the participant moved during breaks, the calibration routine was repeated.

2.3. Stimuli

2.3.1. Stimulus tracking task

The stimulus was a Gaussian blob, presented on a static gray background of 30 cd/m^2. The blob's full-width-at-half-maximum was 0.43 degrees of visual angle. This corresponds to a Goldmann size III stimulus, a commonly used stimulus size in visual function testing.

The stimulus was presented at 5%, 10%, 20%, 40% and 80% Weber contrast, covering a range from near-threshold foveally to suprathreshold peripherally.

2.3.2. Stimulus trajectories

The stimulus moved along a trajectory in one of two tracking modes: smooth pursuit and saccadic pursuit. In the smooth pursuit mode, the stimulus moves continuously, while the direction and velocity of the motion varies in a pseudo-random manner. In this mode, stimulus tracking can, in principle, be done using smooth pursuit eye movements only. We cannot, nor need to, exclude the presence of catch-up saccades. In the saccadic pursuit mode, at pseudo-random intervals, the stimulus additionally makes jumps to new test locations and thus explicitly requires the participant to make saccadic eye movements in order to keep track of the stimulus. As both the timing and amplitude of the stimulus jumps are pseudo-randomized, participants cannot predict when and whereto the stimulus will jump. Therefore, in this mode, the stimulus additionally stimulates the periphery allowing assessment of peripheral contrast sensitivity. A trajectory always started from the center of the screen.

A smooth pursuit trajectory was generated by sampling a velocity value from a zero-centered Gaussian distribution for each frame of the trajectory (40 s at 240 Hz = 9600 frames). This distribution had a standard deviation of 40 deg/s. The values were sampled independently for the horizontal and the vertical components. The resulting velocity array was then low-pass filtered (cutoff = 2 Hz) to minimize jitter and integrated over time to obtain arrays of positional values. The trajectory was generated with the requirement to stay within the boundaries of the screen for its entire duration. A typical resulting smooth trajectory had a root mean square (RMS) stimulus velocity of $\sim$ 10 deg/s with a standard deviation of $\sim$ 5 deg/s.

A saccadic pursuit trajectory was generated by frequently displacing the stimulus in a smooth pursuit trajectory. The time interval between two consecutive displacements was between 1.5 and 2.5 s. An additional requirement was that during a time window of 1 s after a displacement, the stimulus was not allowed to smoothly move towards its starting position before the displacement. Spatially, the displacements were towards one of 56 test locations on a grid. The grid has a resolution that is comparable to a standard grid used for visual field testing, the Humphrey Field Analyzer (HFA, Carl Zeiss Meditec AG, Jena, Germany) 24–2 grid, with a 6 deg spacing (the grid consisted of the 54 locations of a standard HFA 24–2 grid, modified to be symmetrical by inserting two additional test locations on the temporal side). This grid was chosen as it systematically covers the most relevant area of the visual field and would thus potentially also allow for screening for localized reductions in contrast sensitivity. The order in which these locations were sampled was pseudo-randomized, such that in three 40 s saccadic pursuit trajectories all 56 test locations were covered once.

To limit the effect of learning and predictive viewing behavior influencing performance, a total of 10 unique trajectory sets were generated, each consisting of a smooth and three saccadic pursuit trajectories. In addition, a unique practice trajectory set was generated.

2.3.3. Static foveal and peripheral contrast thresholds

To compare performance on the tracking task to static foveal and peripheral Weber contrast thresholds we used a method described in detail elsewhere (Bierings et al., 2018). The contrast thresholds were determined by a 4–2 dB stair-case procedure. Thresholds were measured at the fovea [0 deg, 0 deg] and at four peripheral test locations; [$\pm$ 9 deg, $\pm$ 9 deg]. The stimulus was identical to the one used in the tracking task and presented for 200 ms. The actual assessment was preceded by a practice run.

2.4. Procedure

The experiment started with the static contrast threshold assessment, which took about 5 min. The participant had to fixate the center of the screen and was asked to press a button when a stimulus was perceived. The assessment started at the foveal location followed by the four peripheral test locations. Lastly, the blind spot was probed to assess if fixation was maintained. Next, the eye tracker was calibrated and to familiarize the participant with the tracking task, they were presented with the practice trajectory set. During practice, the tracking stimulus was shown at a level of 100% contrast. The instruction given to the participant was to try and follow the stimulus with their gaze to the best of their abilities, without making any head movements.

Next, for each participant the 10 unique trajectory sets were presented in pseudo-random order such that each trajectory set was presented only once. Each trajectory set started with the smooth pursuit trajectory followed by the three saccadic pursuit trajectories. During each trajectory set, the to-be tracked stimulus was presented at one of the five possible contrast levels, which were assigned first in descending order (80%, 40%, 20%, 10%, 5%) and then in ascending order (5%, 10%, 20%, 40% and 80%). Over all participants, it was ensured that each trajectory set was combined with each contrast level equally often. Prior to presenting these trajectory sets, the participant was alerted to the fact that at lower contrast levels, the stimulus could become very faint or even invisible. We encouraged participants to take breaks between the trajectories whenever needed to maintain concentration and to avoid fatigue. The task was self-paced, i.e., the participant could start the presentation of a trajectory with a mouse click whenever they felt they were ready. A drift correction of the eye tracker calibration was performed prior to presenting each set of trajectories. Overall, including practice, during the tracking task 44 trajectories were presented to the participant, which took about 45 min, including the breaks.

2.5. Analyses

2.5.1. Pre-processing of eye movement data

Eye gaze positions were recorded as horizontal and vertical screen coordinates and analyzed separately. Before feature extraction, the raw gaze data were blink-filtered. Blink periods were identified as velocity spikes (velocity larger than 300 pixels/frame) followed by position plateaus (zero velocity) or missing data. A velocity spike occurs when the eye closes and partially covers the pupil, which the eye-tracker interprets as a rapid upward gaze shift. The blink period was expanded by five samples before and after the defined blink period. The data from the expanded blink period was then interpolated by fitting an autoregressive model using the built-in Matlab function *fillgaps* including ten samples preceding and following each blink period. If the total data loss due to blinks exceeded one-third of the trajectory duration, the entire trajectory was removed from further analyses. After blink filtering, and converting from pixels to degrees of visual angle, the data were low-pass filtered with a first-order Butterworth filter with a half-power frequency of 0.5 Hz.

2.5.2. Tracking performance

After pre-processing, spatio-temporal properties (STP) of eye movements were extracted from gaze data using a method described earlier (Grillini et al., 2020). In total, ten STP can be extracted, four temporal properties, four spatial properties, and two spatio-temporal properties. For brevity and clarity, the analyses described here focus on a single STP, the cosine similarity (Han et al., 2012) of the stimulus and gaze trajectories. Cosine similarity captures overall tracking performance across an entire trajectory in a single value and therefore summarizes overall tracking performance. Evaluating the other STP becomes useful when assessing visual field defects and to distinguish between visual field defects and oculomotor disorders as they show different effects in different STP (Grillini et al., 2020; Soans et al., 2021). The influence of contrast on the other STP can be found in the supplementary material (specifically, supplementary figures A.1 and A.2). In our context, cosine similarity is a measure of the similarity between two vectors: the subsequent stimulus positions and gaze positions during a trajectory. Specifically, it is the normalized inner product of these two vectors and is bound between −1 and 1.

$$CosineSimilarity(A,B) = \frac{A \cdot B}{||A||\,||B||} \tag{1}$$

where A is the vector of stimulus positions and B the vector of gaze positions, $||A||$ the Euclidean norm of vector A, and $||B||$ the Euclidean norm of vector B. The cosine similarity has the useful property of being unaffected by the duration of the trajectory and, since it is computationally inexpensive, is a useful feature to evaluate the performance of a participant in real-time. In healthy participants, it correlates highly with the observation noise (for a full description see supplementary material), a parameter that can be used to measure visual spatial sensitivity (Bonnen et al., 2015).

2.5.3. Detection thresholds

For static contrast threshold assessment, the foveal detection threshold was defined as the log Contrast Sensitivity (logCS) at the foveal test location while the peripheral detection threshold was defined as the average logCS at the four peripheral test locations. Individual logCS values at peripheral test locations that were higher than the mean peripheral logCS plus 2.5 SD of the foveal test location (per age group) were excluded (Chauvenet's criterion; Bierings et al., 2018; Chauvenet, 1906). If the stimulus at one peripheral test location was not seen at all, we assigned a logCS value of −0.2 logCS to that particular test location (Bierings et al., 2018; João et al., 2019). In most cases, the missed test location was the first peripheral test location, most likely due to unfamiliarity with the test despite the practice trial.

For the eye movement data, we ran an analysis to determine at what contrast level the stimulus could be classified as detected. For this, we took the 50% point between chance and ceiling levels as the detection threshold. To determine the chance level, models were created in which the gaze data from a specific trajectory was coupled with the stimulus data from all other trajectories. In this case, the gaze and stimulus trajectories are uncorrelated, except for by chance. These chance models were created for each individual participant. For the ceiling level, stimulus data from each trajectory were coupled with stimulus data from the same trajectory with a delay (which then acted as pseudo-gaze data). In this case, both trajectories should be perfectly correlated except for the added delay. The delay was set at 200 ms, corresponding to the average delay between the stimulus and gaze lag, obtained at the highest contrast level (100%; used during practice) for all participants (we made no distinction between age groups, as at this contrast level, gaze lag did not differ between older and younger participants; see supplementary material for a description). The detection threshold was then set halfway between the median chance level and the median ceiling level for the tracking performance (cosine similarity), separately for the smooth and saccadic pursuit modes. Ceiling and chance levels are shown in Fig. 1.

For smooth pursuit, the ceiling level for tracking performance approached 1 (median: 0.985) and the median chance level was 0.0016 for younger participants and 0.015 for older participants. Therefore, the detection threshold performance (50% point) for smooth pursuit was at approximately 0.5 for all participants. For saccadic pursuit, the ceiling level was further from 1 (median: 0.915), whereas the median chance level was roughly 0 (-0.0009 for younger participants, 0.00006 for older participants). The detection threshold for saccadic pursuit was thus 0.46 for both younger and older participants.

In order to determine the contrast threshold for tracking, the cosine similarity values were linearly interpolated, per participant, and the contrast level at which the tracking performance crossed the detection threshold was calculated (see Fig. 2; the location on the contrast axis where each line crosses the detection threshold was selected as the contrast threshold), separately for the smooth and saccadic pursuit modes. The data were collapsed over the entire trial for smooth pursuit and over all 56 displacements in saccadic pursuit.

2.5.4. Statistical analyses

All statistical analyses were conducted in R (version 4.0.3) using R Studio (version 1.3.959). Linear mixed regression models were performed using the *lmer* function from the *lmerTest* package (3.1–2). Model comparison using the *anova* function from the *stats* package (4.0.3) was done to find the best model structure.

We considered the best model to be the one with the lowest AIC and BIC values. For an overview of all the compared models and corresponding results, see supplementary material. For the best model,

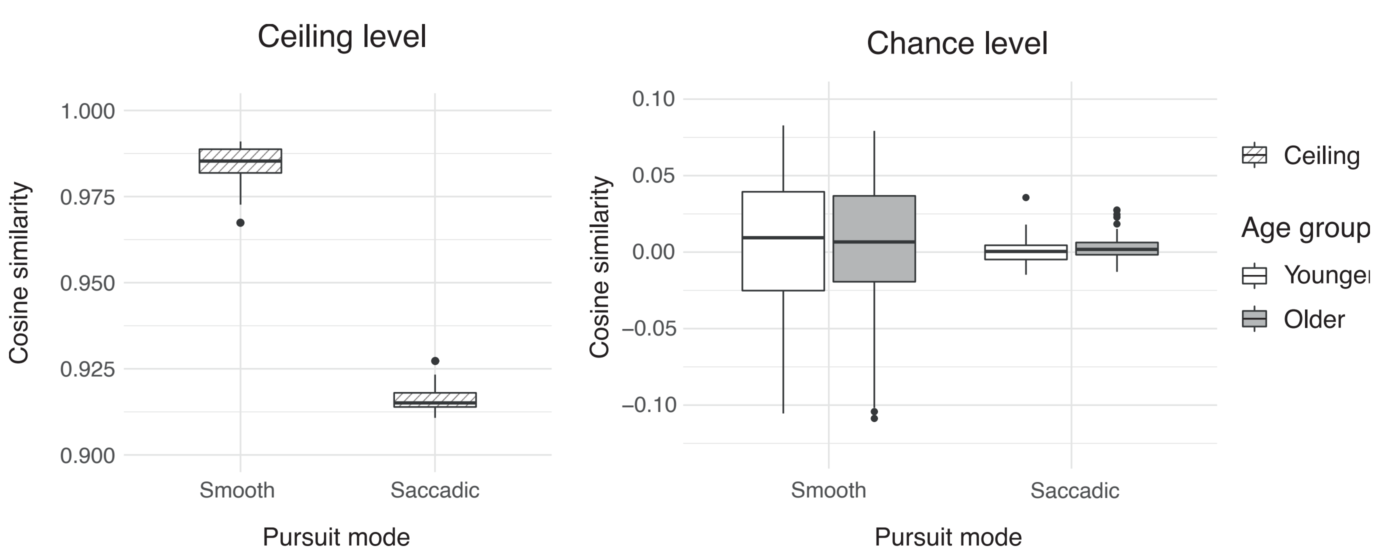

Fig. 1. Left: ceiling level cosine similarity (striped boxes) for smooth and saccadic pursuit modes for all trajectories, averaged over horizontal and vertical data. Right: chance level cosine similarity for smooth and saccadic pursuit modes for younger (white) and older (gray) participants, averaged over horizontal and vertical gaze data. Note that the vertical axes are not the same for the ceiling and chance levels.

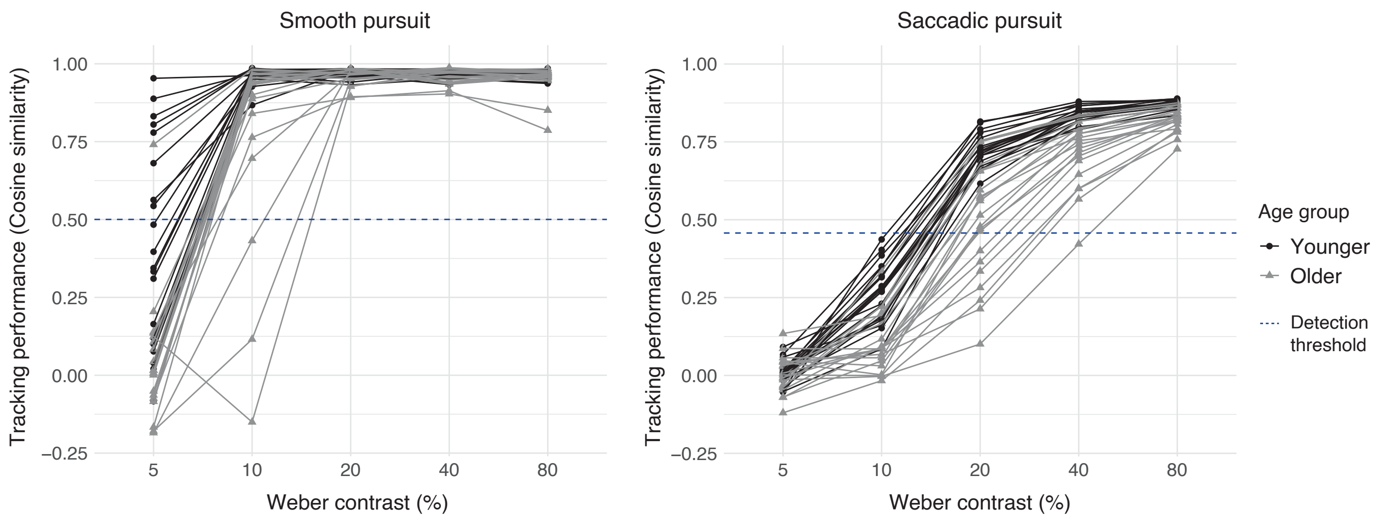

Fig. 2. Tracking performance as a function of stimulus contrast for each individual participant for smooth (left) and saccadic (right) pursuit modes. The results for the younger participants are shown as black dots with connected lines, while those for the older participants are shown as gray triangles with connected lines. The detection threshold is indicated by the dashed blue line (for details, see legend to Fig. 3).

significance of main effects and interactions across factor levels was tested with the *Anova* function from the *car* package (3.0–10), with the used test set to type III Wald chi-square.

To test whether stimulus contrast, pursuit mode, and age affected overall tracking performance (operationalized via cosine similarity), we performed a linear mixed-effects regression with the cosine similarity as the dependent variable. As the effects of contrast and age on the cosine similarity, but also for the other STP, were very similar for horizontal and vertical eye movements (see supplementary Figures A.3 and A.4), we averaged the horizontal and vertical cosine similarities prior to the regression analysis. Regression analyses for horizontal and vertical movements separately can also be found in the supplementary materials. As the patterns of tracking performance as a function of contrast were quite different for smooth and saccadic pursuit (see Results section), separate regression analyses were performed for these. For the smooth and saccadic pursuit modes each, the tested regression models, see supplementary material for details, included *contrast* (factor with five levels with 80% contrast as the reference) and *age group* (with two levels). Each contrast level was presented twice reflected in the models as the factor *presentation* with two levels (first and second presentation). *Participant* was included as a random intercept. No random slopes were included as the models did not converge. The regression models violated the assumptions of normality and homoscedasticity of the residuals (see supplementary material, Figure A.7). To confirm the model estimates we therefore used bootstrapping to calculate the confidence intervals (using the function *confint* from the *stats* package) with 1000 replications. Estimates for which the confidence intervals did not straddle zero were considered reliably different from zero.

An additional linear mixed-effects regression was performed on the static contrast thresholds data to test whether logCS differed between the younger and older participants and between the foveal and peripheral test locations. The tested models, see supplementary material, included *eccentricity* (with two levels: fovea, periphery) and *age group* as well as a random intercept for *participant*. Again, no random slope was included as the models did not converge.

Lastly, we performed a correlation analysis between the measured static Weber contrast threshold (the inverse of the CS) and the threshold for the stimulus tracking. As the stimulus in smooth pursuit mostly stimulates central vision, in the analysis, the smooth pursuit detection thresholds were correlated with the static foveal contrast thresholds. In the saccadic pursuit the periphery is additionally stimulated, thus the saccadic pursuit detection thresholds were correlated with the static peripheral contrast thresholds. As the data were not normally distributed, a Spearman correlation test was performed using the *cor.test* function from the *stats* package (Best & Roberts, 1975).

3. Results

3.1. Continuous tracking performance as a function of stimulus contrast

Figs. 2 and 3 show individual (Fig. 2) and group-averaged (Fig. 3) tracking performance (expressed in terms of cosine similarity) as a function of stimulus contrast for the smooth (left) and saccadic (right) pursuit modes, respectively. For both the smooth and saccadic pursuit modes, tracking performance increased with increasing stimulus contrast. To reach the same level of performance, the older participants needed more contrast than the younger participants.

For the smooth pursuit mode (left panels in Figs. 2 and 3), at 5% contrast, performance for most younger (12/20) and all but one of the older participants was below threshold. From 10% contrast onwards, all younger participants performed above the detection threshold and nearly all performed at ceiling level. Most older participants performed above the detection threshold at 10% contrast, but only from 20% contrast onwards all of them performed above threshold and most of them reached performance at ceiling level. The best regression model (see Table A.1 in supplementary material) included the factors *contrast* (with five levels with 80% contrast as the reference) and *age group* (with two levels, younger and older age) as main effects and in interaction, and *presentation* (repetition of contrast levels; two levels, first and second presentation) as main effect only as well as *participant* as a random intercept: Cosine similarity ~ contrast * age group + presentation + (1| participant). The model showed a significant main effect of *contrast* ($Chi^2(4) = 384.77$, $p < 0.001$), indicating increased tracking performance at higher contrast, and a significant interaction between *contrast* and *age group* ($Chi^2(4) = 81.89$, $p < 0.001$), indicating different effects of contrast in the two age groups. The main effects of *age group* ($Chi^2(1) = 0.18$, $p = 0.67$) and of presentation ($Chi^2(1) = 0.07$, $p = 0.79$) were not significant. Confidence intervals from the bootstrapping procedure confirmed these findings. A summary can be found in Table 1. The bootstrapping showed that for the younger participants, compared to tracking performance at 80% contrast, performance was reliably lower at 5% contrast but not at 10%, 20%, and 40% contrast. This indicates that as a group the younger participants reached performance at ceiling level around 10% contrast. For the older participants, however, tracking

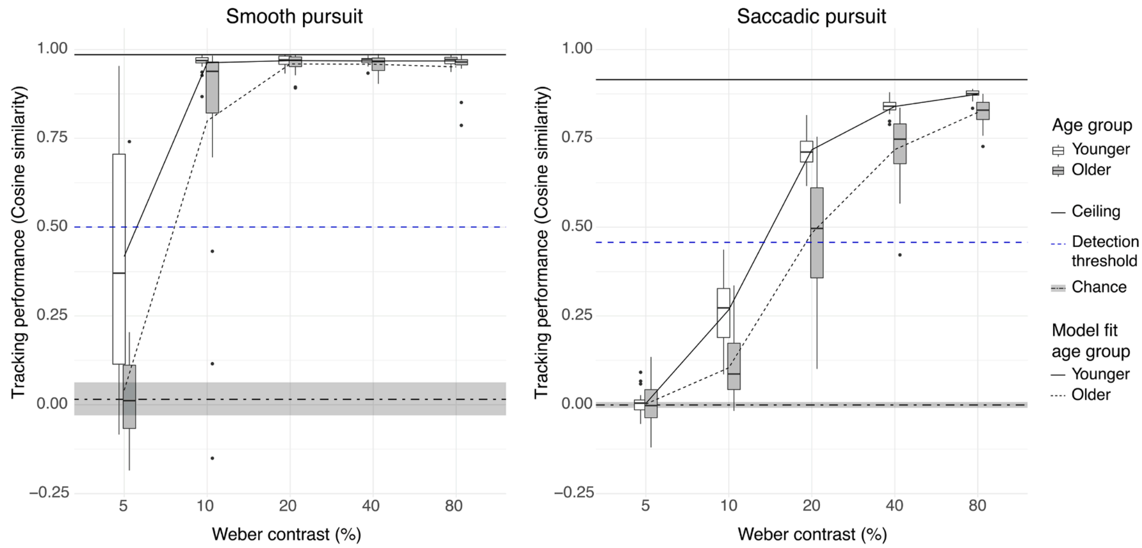

Fig. 3. Tracking performance for smooth (left) and saccadic (right) pursuit modes for younger (white) and older (gray) participants at the different stimulus contrast levels. The solid black curve shows the regression model fit for younger participants while the dashed black curve shows the regression model fit for older participants. The black horizontal line (at the top) shows the ceiling level. The dash-dotted horizontal black line (at the bottom) shows the median chance level with the shaded area indicating the standard deviation. The dashed horizontal blue line indicates threshold tracking performance above which the stimulus is considered to be seen, and is the midpoint between the chance and ceiling levels.

Table 1
Model and bootstrap summaries for the smooth and saccadic pursuit mode results. The intercept represents the estimate for younger participants at the first presentation of 80% Weber contrast. All other parameters are with respect to this intercept. The 95th percentile confidence intervals (CI) were calculated using bootstrapping. If the CI does not straddle 0, it indicates the estimate is reliably different from 0. These estimates are printed in bold typeface.

	Smooth pursuit mode	**Saccadic pursuit mode**
	Estimate [2.5%, 97.5%]	Estimate [2.5%, 97.5%]
Intercept (80% contrast, younger age)	**0.97 [0.91, 1.02]**	**0.87 [0.84, 0.70]**
40% contrast, younger age	0.00 [−0.07, 0.07]	**−0.03 [−0.06, 0.00]**
20% contrast, younger age	0.00 [−0.07, 0.07]	**−0.16 [−0.19, −0.13]**
10% contrast, younger age	−0.01 [−0.08, 0.07]	**−0.61 [−0.4, −0.58]**
5% contrast, younger age	**−0.55 [−0.62, −0.48]**	**−0.87 [−0.90, −0.84]**
Older age	−0.02 [−0.09, 0.06]	**−0.05 [−0.10, 0.00]**
Second presentation	0.00 [−0.04, 0.03]	0.00 [−0.01, 0.02]
40% contrast, older age	0.01 [−0.09, 0.11]	**−0.07 [−0.12, −0.03]**
20% contrast, older age	0.01 [−0.09, 0.11]	**−0.18 [−0.23, −0.14]**
10% contrast, older age	**−0.15 [−0.25, 0.04]**	**−0.11 [−0.16, −0.07]**
5% contrast, older age	**−0.36 [−0.46, −0.26]**	**0.05 [0.01, 0.09]**

performance was lower at both 5% and 10% contrast compared to that at 80% contrast, indicating that, as a group, they reached performance at ceiling level around 20% contrast.

For the saccadic pursuit mode (right panels in Figs. 2 and 3), at 5% and 10% contrast, all the younger and older participants performed below threshold. From 20% contrast onwards, all younger participants performed above the detection threshold. At 20% contrast, the performance of most of the older participants (13/20) exceeded the detection threshold, while at 40% contrast all but one of them performed above threshold. At 80% contrast, the tracking performance of both groups approached the ceiling level. Again, the best regression model (see Table A.1 in supplementary material) included *contrast* and *age group* as main effects and in interaction, *presentation* as main effect only and *participant* as a random intercept: Cosine similarity ~ contrast * age group + presentation + (1|participant). The model showed significant main effects of *contrast* ($Chi^2(4) = 8868.81$, $p < 0.001$), indicating increased performance at higher contrast levels, and *age group* ($Chi^2(1) = 34.46$, $p < 0.001$), indicating overall higher performance for the younger participants. In addition, there was a significant interaction between *contrast* and *age group* ($Chi^2(4) = 134.38$, $p < 0.001$), indicating different effects of contrast in the two age groups. *Presentation* was not significant ($Chi^2(1) = 0.49$, $p = 0.49$), indicating stable performance throughout the experiment. Bootstrapping (see Table 1) showed that for both age groups, tracking performance differed reliably at all contrast levels, which may be taken as an indication that tracking performance had not yet plateaued.

3.2. Static contrast thresholds compared to smooth and saccadic pursuit detection thresholds

Fig. 4 shows the static foveal (left) and peripheral (right) thresholds for each participant versus the smooth and saccadic pursuit detection thresholds, respectively, and the correlation between both measures. The mean foveal contrast threshold was 4% (range 3–6%) for the younger participants and 5% (range 4–8%) for the older participants. The mean of the averaged contrast threshold of the peripheral test locations was 15% (range 13–18%) and 20% (range 12–32%) for the younger and older participants, respectively. The older participants had higher foveal and peripheral Weber contrast thresholds compared to the younger participants (the difference was approximately 0.11 log units in

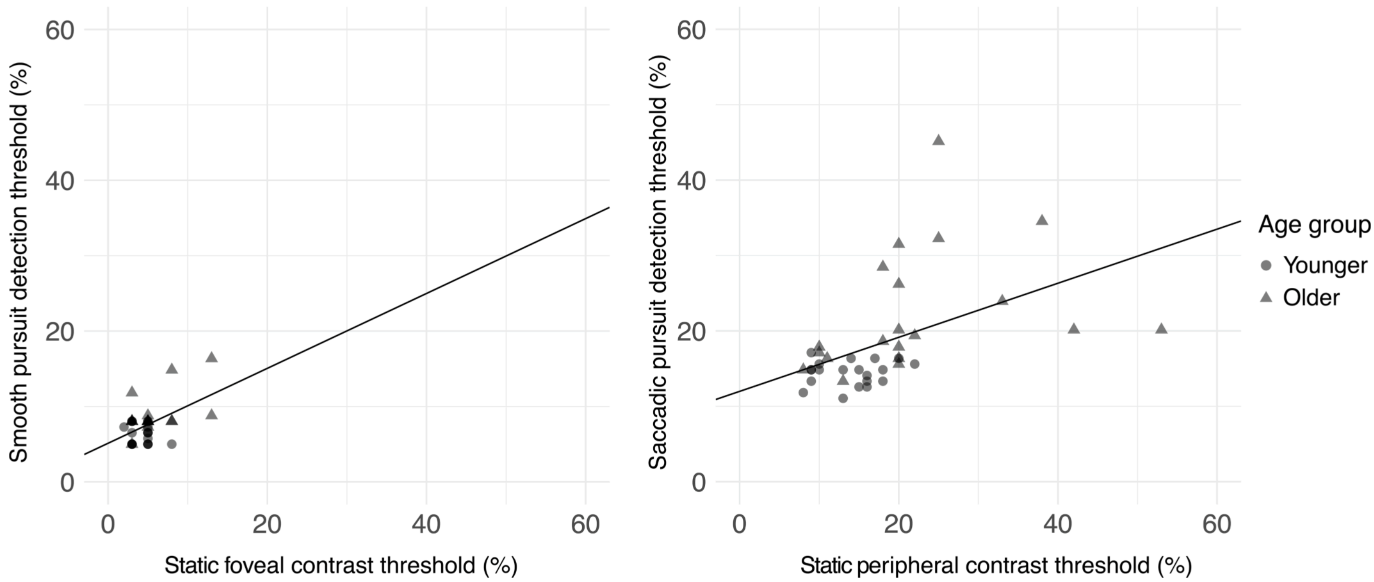

Fig. 4. Scatterplots for the static foveal contrast thresholds vs. the smooth pursuit mode thresholds (left) and for the static peripheral contrast thresholds vs. the saccadic pursuit mode thresholds (right). All thresholds are Weber contrasts. The data for the younger participants are shown as semi-transparent circles, while data from older participants are shown as semi-transparent triangles. Please note that threshold values overlap for some participants, especially in the smooth pursuit mode and static foveal thresholds. These overlapping values are plotted on top of each other, indicated by less transparent (darker) circles/triangles. The regression lines in the panels give an indication of the correlation between both measures. The correlation between the peripheral and the saccadic pursuit thresholds was significant (rho = 0.64, $p < 0.001$), while the correlation between the foveal and the smooth pursuit thresholds was not (rho = 0.29, p = 0.07).

the fovea and approximately 0.18 log units in the periphery). The best regression model (see Table A.2 in supplementary material) included main effects for *eccentricity* (with two levels: fovea and periphery) and *age group* as well as a random intercept for *participant*: logCS ∼ eccentricity + age group + (1|participant). The model confirmed that contrast thresholds were higher (i.e., lower logCS) in the periphery than in the fovea (estimate = -0.54, std. error = 0.03, t = -17.30, $p < 0.001$) and that the younger participants had a lower contrast threshold (higher logCS) than the older participants (estimate = 0.14, std. error = 0.04, t = 3.28, p = 0.001).

The smooth pursuit and foveal contrast thresholds did not correlate significantly (rho = 0.29, S = 7557.3, p = 0.07). There was a moderate positive correlation between the saccadic pursuit and peripheral contrast thresholds (rho = 0.64, S = 3870.9, $p < 0.001$), indicating that the contrast thresholds measured with the more conventional static contrast threshold assessment are predictive of the detection thresholds measured by means of continuous saccadic pursuit tracking.

4. Discussion

Our main finding is that continuous tracking performance strongly depends on stimulus contrast, age, and stimulus movement (pursuit mode). With increasing stimulus contrast, tracking performance increases in both the younger and older participants. The influence of contrast differs between the two pursuit modes, with participants requiring about twice as much contrast to perform the saccadic pursuit compared to the smooth pursuit mode. Age appears to be an important factor as well, since older participants required about twice as much contrast, compared to the younger participants, to reach threshold performance in both the smooth and saccadic pursuit modes. The detection thresholds determined by means of saccadic pursuit tracking correlate moderately with the peripheral static contrast thresholds. Below, we discuss our results and their implications in more detail.

4.1. Tracking performance depends on stimulus contrast

We found that tracking performance depended on the stimulus contrast, with an increasing level of performance at higher contrasts for both the smooth and saccadic pursuit modes in both younger and older participants. This important role of contrast is in agreement with Mazumdar et al. (2019) who investigated the interaction of the participant's age, sex, stimulus intensity, and eccentricity on saccadic reaction time (SRT) behavior in eye movement perimetry (EMP). Using four stimulus contrasts (26%, 41%, 64%, and 82%), they showed that SRTs became shorter with increasing stimulus contrast and decreasing eccentricity.

In our study, for smooth pursuit mode, both younger and older participants display a complete psychometric curve over a relatively narrow contrast range, with performance at chance level for 5% contrast, and reaching performance at ceiling level at 10% and 20% contrast, respectively. For saccadic pursuit mode, both age groups performed at chance level at 5% contrast and then showed an increase in tracking performance with increasing contrast. At 80% contrast, the younger participants' performance approached the empirical ceiling level (0.92) with a median cosine similarity of 0.88. With a median cosine similarity of 0.83 at 80% contrast, the older participants did not yet reach this ceiling level performance. Nevertheless, for both age groups and both pursuit modes, the contrast levels chosen were in the midrange of the psychometric curve, where the sensitivity for detecting impairments is greatest. Furthermore, we found that for the smooth pursuit mode, the threshold contrast was between 5% and 10% for the younger participants and between 10% and 20% for the older participants. For the saccadic pursuit mode, the threshold contrast was between 10% and 20% for the younger participants and between 20% and 40% for older participants. Therefore, testing at more stimulus contrasts between 5% and 40% contrast could provide more precise threshold estimates.

4.2. Obtaining similar tracking performance requires more contrast in older compared to younger participants

For both smooth and saccadic pursuit modes, on average, stimulus contrast needed to be approximately doubled for older participants compared to younger participants to cross the detection threshold. This difference between the older and the younger participants is in agreement with the overall age-related reduction in contrast sensitivity (Haegerstrom-Portnoy et al., 1999) and the age-related reduction in contrast sensitivity in the normal visual field as measured by SAP (Heijl et al., 1987; Katz & Sommer, 1986). Reduced contrast sensitivity with increasing age is also found in semi-automated kinetic perimetry (Vonthein et al., 2007). Vonthein and colleagues (2007) showed an increasingly greater decline in the area of visible stimuli while maintaining fixation (isopter area) with a reduction in stimulus size and stimulus luminance with older ages (Vonthein et al., 2007). Several studies that map the visual field by EMP have investigated the effect of age on the SRT of eye movements in healthy controls and have shown that age has a significant effect on SRT. They found a significant delay in SRT with increasing age (Mazumdar et al., 2019; Munoz et al., 1998). The age-related reduction in performance for saccadic pursuit suggests that age-specific normative databases will be required to classify individual performance as impaired in the future. This requirement is in agreement with the existence of age-specific normative databases in conventional static (Katz & Sommer, 1986) and kinetic perimetry (Vonthein et al., 2007) and with EMP (Kadavath Meethal et al., 2018; Mazumdar et al., 2014, 2019, 2020; Pel et al., 2013).

4.3. *Better tracking performance in the smooth compared to the saccadic pursuit mode.*

Overall, at the same level of contrast, we found better performance during the smooth pursuit mode compared to the saccadic pursuit mode. This difference is not surprising since in the saccadic pursuit mode, the participant is forced to use their peripheral vision and thus needs to make saccadic eye movements to keep track of the stimulus, while this is not necessarily required in the smooth pursuit mode. Visual sensitivity decreases towards the periphery, thus straightforwardly explaining this difference in performance, at least for controls. However, in future applications of our screening task in patients this may be different as various ophthalmic and neurological disorders may differentially affect central and peripheral contrast sensitivity and thus result in deviations from this expected pattern.

4.4. *Agreement between saccadic tracking performance and static contrast thresholds*

Our results showed a positive correlation between the saccadic pursuit detection thresholds and the peripheral contrast threshold, indicating that the static contrast thresholds are predictive of the performance on saccadic pursuit tracking (as operationalized via cosine similarity). However, the smooth pursuit detection thresholds and static foveal contrast thresholds did not show such a correlation. As all participants had good visual acuity, their foveal contrast detection thresholds were low in both the static assessment and the smooth pursuit tracking, and the variance in these thresholds across participants was also low. This lack of variance makes it harder to detect a possible correlation. In patient populations, larger differences in foveal sensitivity and thus more variance and higher correlations between the detection thresholds based on smooth pursuit tracking and static contrast threshold assessment can be expected.

4.5. *Limitations and future directions*

4.5.1. *The cosine similarity is a good overall measure of performance, but other spatio-temporal properties are important when considering patient populations*

In the current study, we used the cosine similarity as our summary measure for tracking performance. Based on a cross-correlogram analysis of the stimulus trajectories and eye movements (Mulligan et al., 2013), we can derive various additional spatio-temporal properties (STP, for a description, see supplementary material). These STP, for example, quantify the positional error and the lag between the gaze and the stimulus and provide a detailed impression of the integrity of various visual and oculomotor functions. In their simulation study, Grillini et al. (2018) showed that the STP of healthy participants changed when visual field defects (VFD) were simulated, and these changes were specific to the type of VFD (central scotoma or peripheral scotoma). However, estimating several STP involve model fitting, i.e. modeling the raw eye movement cross-correlogram with a gaussian function. At 5% contrast (and in some cases also at 10% contrast), the gaze and stimulus were highly uncorrelated (indicated by a low cosine similarity), such that it was impossible to get a good model fit. Consequently, the resulting STP (such as the amplitude and width) were highly variable and unreliable. In addition, all other STP showed a very similar relationship to contrast, namely better values (higher or lower depending on the specific STP) for higher contrast and generally better values for younger compared to older participants. Nevertheless, evaluating other STP will be important when examining specific (neuro-)ophthalmic disorders as these could lead to changes in specific STP (Grillini et al., 2018, 2020; Soans et al., 2021). In individuals with hemianopia for example, a marked positional bias can be expected due to the loss of a full hemifield (see Grillini et al. (2018) for STP results in simulated visual field defects). Grillini et al. (2020) showed, at the group level, deviations in specific temporal and spatial STP in individuals with Multiple Sclerosis primarily in saccadic pursuit, whereas in individuals with Parkinson's Disease nearly all evaluated STP showed deviations in both pursuit modes (Grillini et al., 2020). Thus, when investigating patient populations, it is imperative to take all STP for smooth and saccadic pursuit into account, as patients may differ from controls on some but not all STP. The same reasoning applies to analyzing horizontal and vertical components separately. In controls, it is not expected that horizontal and vertical components differ, which is why we averaged them. However, impaired visual function could affect horizontal and vertical components differently.

4.5.2. *Dependency of tracking performance on eccentricity should be considered*

Previous studies using static contrast threshold assessments (Heijl et al., 1987; Katz & Sommer, 1986) and EMP studies showed that the age-related reduction in sensitivity is larger at higher eccentricity. Our present performance indicator, cosine similarity, measures the overall agreement between the eye and stimulus movements in a trial. Consequently, information on tracking performance at specific eccentricities is lost. Because of this, we can only speculate that in our task, the age-related sensitivity reduction will increase with eccentricity. One way to study this dependence in our task would be by separately estimating the sensitivity at each grid location. Subsequently, these estimates could be compared between different age groups. Based on these estimates, it would also be possible to compare the location-wise correspondence in sensitivities determined using conventional static contrast threshold assessment and continuous stimulus tracking. However, we have not done this for the present paper, as estimating sensitivity at each grid location requires a rather different type of analysis than the current one (Grillini et al., 2021).

4.5.3. *Use of eccentricity dependent stimulus contrast could improve contrast threshold assessments*

We found that, in order to reach the detection threshold, a lower contrast was required in the smooth pursuit compared to the saccadic pursuit mode. Therefore, it would be logical to proceed the development of this technique with different contrasts for the different pursuit modes, in order to facilitate the precise measurement of contrast thresholds at the different test locations. Moreover, stimuli could be presented at eccentricity-dependent contrast and stimulus size (Bedggood et al., 2021; Heijl et al., 1987; Redmond et al., 2010; Rountree et al., 2018). Albeit attractive, this approach requires information on momentary gaze location and would thus require on-line analysis. Further research should uncover if the performance of this - potentially superior -

approach would surpass the robust simplicity of a stimulus that is fixed in contrast (except for saccadic versus smooth pursuit mode) and size.

4.5.4. Bootstrapping confirms results from linear regression models

While our regression models fit the data well, the assumptions for mixed linear regression were violated and the model estimates and their significance may not be correct. The likely reason for the violation of the assumptions is the fact that the dependent variable (cosine similarity) is bound between −1 and + 1, where generally in linear regression with Gaussian distributions the dependent variable is unbounded. Because of these assumption violations, we performed bootstrapping to confirm the significance of model estimates. The bootstrapping did not change the result from the regression models, and other studies have shown that linear mixed-effects models are fairly robust to violation of assumptions (Arnau et al., 2013; McCulloch & Neuhaus, 2011; Schielzeth et al., 2020), reassuring that the models we built are robust to these violations.

4.6. Clinical application

As already alluded to above, based on a rather different type of analyses, in principle one could obtain sensitivity estimates for each separate grid location. If these estimates could be obtained reliably and with sufficient sensitivity, it would potentially turn our present continuous tracking approach into an alternative form of perimetry. We have several indications that this is feasible. First, in a series of proof-of-principle studies, we have shown that we can detect and characterize both simulated and actual visual field defects (Grillini et al., 2018, 2021; Soans et al., 2021). In addition, based on tracking data acquired during task performance with gaze-contingent simulated scotoma, Grillini et al. (2021) showed that it is possible to reconstruct visual field maps, including the location of a scotoma. Our present finding that static contrast thresholds correlate with saccadic pursuit performance is a further indication that delineation of peripheral scotoma might be feasible. However, thus far, the simulated scotomas were absolute and the patient populations in these studies were relatively small. Therefore, future studies with more realistic and relative (less deep) simulated scotomas and in larger patient populations are needed to examine whether continuous stimulus tracking can be applied towards accurate scotoma delineation. To be able to do so, stimuli may need to be shown at multiple, clinically meaningful, contrasts. With our current setup, which had a minimum screen luminance of 0.4 cd/m^2, a maximum screen luminance of 358 cd/m^2, and using a background luminance of 30 cd/m^2, we were able to present stimulus contrasts from 2.7% up to 1100% (range of 2.6 log). Our background luminance was brighter than the 10 cd/m^2 background luminance used in conventional SAP. We chose a luminance of 30 cd/m^2 since at this level, the contrast sensitivity thresholds saturate for both controls and glaucoma patients (Bierings et al., 2018). This is in agreement with Weber's law which states that contrast sensitivity is constant at higher background luminances (Duke-Elder & Weale, 1968). At 10 cd/m^2, Weber's law may not always hold and pupil size could affect contrast sensitivity (Anderson & Patella, 1999).

Taking the contrast thresholds for smooth and saccadic pursuit modes as age-normal thresholds, we can cover contrasts over a range of 1.7 log units (20–1100% contrast) for the older participants and 2.0 log units for the younger participants (10–1100% contrast), respectively. This is well within the range described as clinically meaningful by Gardiner & Mansberger (2016). Here, we tested only young (18–29 years of age) and older (51–77 years of age) healthy participants to investigate whether there is an effect of age on the tracking performance. As a next step, to create an age-specific normative database for the tracking task, we should additionally test healthy participants in the age categories 30–39 and 40–49. With this database, we would be able to classify individual performance as impaired. When using these age-normal thresholds as baseline stimulus contrasts, even slight reductions in contrast sensitivity should lead to a substantial drop in tracking performance. Such age-normal thresholds could allow for the detection of relative and early visual field defects (Hodapp et al., 1993) which would be a major advance over the results of our previous proof-of-principle studies. In addition, to further establish the potential of continuous visual tracking to differentiate between different types and severities of visual field defects (Hodapp et al., 1993) and underlying pathologies, clinical studies with participant groups with different types and severities of visual field defects should be performed. At present, it is difficult to predict how tracking performance would be affected when the target enters a scotoma (Bennett et al., 2010; Bennett & Barnes, 2004; Orban de Xivry et al., 2006), or whether target direction influences performance (González et al., 2018) or how non-visual factors such as attention could influence the tracking behavior. Especially for central visual field loss, characterizing the shape and severity of the defect could be challenging due to for example the presence of a preferred retinal locus (PRL) (Crossland et al., 2011), experiencing perceptual filling-in of the scotoma (Zur & Ullman, 2003), or whether or not the entire target is visible (González et al., 2019). Therefore, how central visual field defects exactly affect tracking performance and thus the ability to the detect a scotoma using continuous stimulus tracking paradigms needs further research.

Although continuous stimulus tracking is designed to be less affected by calibration-related issues compared to static eye movement perimetry (Grillini et al., 2018, 2021), future clinical application would profit from an eye-tracker that does not require extensive calibration. Eye-trackers that meet this requirement are becoming available (Barsingerhorn et al., 2018; Pupil Labs GmbH, Berlin).

4.7. Conclusion

The present study shows that tracking performance increases with increased contrast for both pursuit modes and that, with age, a higher contrast is needed to maintain performance. In addition, the pursuit mode affects tracking performance, with better tracking performance for the smooth compared to the saccadic pursuit mode. With this study, we have determined an optimal range of contrasts required to present stimuli in the midrange of the psychometric curve where even slight reductions in contrast sensitivity should lead to a decrease in tracking performance, thus allowing detecting subtle impairments. These findings suggest various opportunities for the development of continuous visual stimulus tracking towards a clinically viable visual function screening tool. Notably, continuous visual stimulus tracking as done here enables to capture more aspects of the functioning of the visual system than trial-based approaches, such as systematic delays in the tracking. Such information can be relevant for characterizing neurological diseases (Grillini et al., 2020). The present findings may also help to optimize this application of continuous visual stimulus tracking.

Open Practices Statement: The data are available at the DataverseNL repository via https://doi.org/10.34894/CHFEB4, while materials (experimental code) are available at request from the corresponding author. The experiments were not pre-registered.

Funding: This project was funded by the Vision Foundation, The Netherlands and ZonMw, programme Expertisefunctie Zintuiglijk Gehandicapten, grant number 637005001, and by Uitzicht grant number UZ2019-20 (via funds provided by the Algemene Nederlandse Vereniging ter Voorkoming van Blindheid, Oogfonds, Stichting blindenpenning, Landelijke Stichting voor Blinden en Slechtzienden). The funding organizations had no role in designing, conducting, analyzing, or publishing this research.

Declaration of Competing Interest

AG is the majority shareholder of Reperio B.V., a private company that develops ophthalmic and neurological tests based on eye movements. ECP is an employee of Reperio B.V.. AG and RJR are listed as inventors on the patent application WO2021096361A1 (Grillini, Renken, & Hernández-García, 2021). Method, system, and computer program product for mapping a visual field) on which the method used in this manuscript is partially based. Reperio B.V. had no control over the design, conduct, analysis, or publication of this research, nor did it contribute financially towards it.

Data availability

The data are available at the DataverseNL repository via https://doi.org/10.34894/CHFEB4, while materials (experimental code) are available at request from the corresponding author.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.visres.2023.108188.

References

Anderson, D. R., & Patella, V. M. (1999). *Automated static perimetry (2nd ed)*. Mosby.

Arnau, J., Bendayan, R., Blanca, M. J., & Bono, R. (2013). The effect of skewness and kurtosis on the robustness of linear mixed models. *Behavior Research Methods, 45*(3), 873–879. https://doi.org/10.3758/s13428-012-0306-x

Barsingerhorn, A. D., Boonstra, F. N., & Goossens, J. (2018). Development and validation of a high-speed stereoscopic eyetracker. *Behavior Research Methods, 50*(6), 2480–2497. https://doi.org/10.3758/s13428-018-1026-7

Bedggood, P., Prea, S. M., Kong, Y. X. G., & Vingrys, A. J. (2021). Scaling the size of perimetric stimuli reduces variability and returns constant thresholds across the visual field. *Journal of Vision, 21*(11), 2. https://doi.org/10.1167/jov.21.11.2

Bengtsson, B., Heijl, A., & Olsson, J. (1998). Evaluation of a new threshold visual field strategy, SITA, in normal subjects: Evaluation of a new threshold visual field strategy, SITA, in normal subjects. *Acta Ophthalmologica Scandinavica, 76*(2), 165–169. https://doi.org/10.1034/j.1600-0420.1998.760208.x

Bennett, S. J., & Barnes, G. R. (2004). Predictive smooth ocular pursuit during the transient disappearance of a visual target. *Journal of Neurophysiology, 92*(1), 578–590. https://doi.org/10.1152/jn.01188.2003

Bennett, S. J., Orban de Xivry, J.-J., Lefèvre, P., & Barnes, G. R. (2010). Oculomotor prediction of accelerative target motion during occlusion: Long-term and short-term effects. *Experimental Brain Research, 204*(4), 493–504. https://doi.org/10.1007/s00221-010-2313-4

Best, D. J., & Roberts, D. E. (1975). Algorithm AS 89: The upper tail probabilities of Spearman's rho. *Applied Statistics, 24*(3), 377. https://doi.org/10.2307/2347111

Bierings, R. A. J. M., de Boer, M. H., & Jansonius, N. M. (2018). Visual performance as a function of luminance in glaucoma: The De Vries-Rose, Weber's, and Ferry-Porter's Law. *Investigative Opthalmology & Visual Science, 59*(8), 3416. https://doi.org/10.1167/iovs.17-22497

Bonnen, K., Burge, J., Yates, J., Pillow, J., & Cormack, L. K. (2015). Continuous psychophysics: Target-tracking to measure visual sensitivity. *Journal of Vision, 15*(3), 14. https://doi.org/10.1167/15.3.14

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Chauvenet, W. (1906). *A Manual of Spherical and Practical Astronomy: Vol. Volume 2* (5th ed.). J.B. Lippincott Company.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink toolbox: Eye tracking with MATLAB and the psychophysics toolbox. *Behavior Research Methods, Instruments, & Computers, 34*(4), 613–617. https://doi.org/10.3758/BF03195489

Crossland, M. D., Engel, S. A., & Legge, G. E. (2011). The Preferred retinal locus in macular disease: Toward A consensus definition. *Retina, 31*(10), 2109–2114. https://doi.org/10.1097/IAE.0b013e31820d3fba

Duke-Elder, S., & Weale, R. A. (1968). The sensation of light. In S. Duke-Elder (Ed.), *System of Ophthalmology* (Vol. IV, p. 583). St. Louis: CV Mosby.

Gardiner, S. K., & Mansberger, S. L. (2016). Effect of restricting perimetry testing algorithms to reliable sensitivities on test-retest variability. *Investigative Opthalmology & Visual Science, 57*(13), 5631. https://doi.org/10.1167/iovs.16-20053

González, E. G., Liu, H., Tarita-Nistor, L., Mandelcorn, E., & Mandelcorn, M. (2019). Smooth pursuit of amodally completed images. *Experimental Eye Research, 183*, 3–8. https://doi.org/10.1016/j.exer.2018.07.015

González, E. G., Tarita-Nistor, L., Mandelcorn, E., Mandelcorn, M., & Steinbach, M. J. (2018). Mechanisms of image stabilization in central vision loss: Smooth pursuit. *Optometry and Vision Science, 95*(1), 60–69. https://doi.org/10.1097/OPX.0000000000001161

Grillini, A., Hernández-García, A., Renken, R. J., Demaria, G., & Cornelissen, F. W. (2021). Computational methods for continuous eye-tracking perimetry based on spatio-temporal integration and a deep recurrent neural network. *Frontiers in Neuroscience, 15*, Article 650540. https://doi.org/10.3389/fnins.2021.650540

Grillini, A., Renken, R. J., & Hernandez Garcia, A. (2021). Method, system and computer program product for mapping a visual field. (Patent No. WO2021096361).

Grillini, A., Ombelet, D., Soans, R. S., & Cornelissen, F. W. (2018). Towards using the spatio-temporal properties of eye movements to classify visual field defects. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications* (pp. 1–5). https://doi.org/10.1145/3204493.3204590

Grillini, A., Renken, R. J., Vrijling, A. C. L., Heutink, J., & Cornelissen, F. W. (2020). Eye movement evaluation in multiple sclerosis and Parkinson's disease using a standardized oculomotor and neuro-ophthalmic disorder assessment (SONDA). *Frontiers in Neurology, 11*, 971. https://doi.org/10.3389/fneur.2020.00971

Haegerstrom-Portnoy, G., Schneck, M. E., & Brabyn, J. A. (1999). Seeing into old age: Vision function beyond acuity. *Optometry and Vision Science: Official Publication of the American Academy of Optometry, 76*(3), 141–158. https://doi.org/10.1097/00006324-199903000-00014

Han, J., Kamber, M., & Pei, J. (2012). Getting to Know Your Data. In *Data Mining* (pp. 39–82). Elsevier. https://doi.org/10.1016/B978-0-12-381479-1.00002-2.

Heijl, A., Lindgren, G., & Olsson, J. (1987). Normal variability of static perimetric threshold values across the central visual field. *Archives of Ophthalmology, 105*(11), 1544–1549. https://doi.org/10.1001/archopht.1987.01060110090039

Hodapp, E., Parrish, R. K., & Anderson, D. R. (1993). *Clinical decisions in glaucoma*. Mosby.

João, C. A. R., Scanferla, L., & Jansonius, N. M. (2019). Retinal contrast gain control and temporal modulation sensitivity across the visual field in glaucoma at photopic and mesopic light conditions. *Investigative Opthalmology & Visual Science, 60*(13), 4270. https://doi.org/10.1167/iovs.19-27123

Jones, P. R., Smith, N. D., Bi, W., & Crabb, D. P. (2019). Portable perimetry using eye-tracking on a tablet computer—A feasibility assessment. *Translational Vision Science & Technology, 8*(1), 17. https://doi.org/10.1167/tvst.8.1.17

Kadavath Meethal, N. S., Mazumdar, D., Asokan, R., Panday, M., van der Steen, J., Vermeer, K. A., … Pel, J. J. M. (2018). Development of a test grid using Eye Movement Perimetry for screening glaucomatous visual field defects. *Graefe's Archive for Clinical and Experimental Ophthalmology, 256*(2), 371–379. https://doi.org/10.1007/s00417-017-3872-x

Katz, J., & Sommer, A. (1986). Asymmetry and variation in the normal Hill of vision. *Archives of Ophthalmology, 104*(1), 65–68. https://doi.org/10.1001/archopht.1986.01050130075023

Kleiner, M., Brainard, D., & Pelli, D. (2007). *What's new in Psychtoolbox-3?*

Mazumdar, D., Meethal, N. S. K., Panday, M., Asokan, R., Thepass, G., George, R. J., … Pel, J. J. M. (2019). Effect of age, sex, stimulus intensity, and eccentricity on saccadic reaction time in eye movement perimetry. *Translational Vision Science & Technology, 8*(4), 13. https://doi.org/10.1167/tvst.8.4.13

Mazumdar, D., Pel, J. J. M., Kadavath Meethal, N. S., Asokan, R., Panday, M., Steen, J. v. d., & George, R. (2020). Visual Field Plots: A Comparison Study Between Standard Automated Perimetry and Eye Movement Perimetry. *Journal of Glaucoma, 29*(5), 351–361. https://doi.org/10.1097/IJG.0000000000001477.

Mazumdar, D., Pel, J., Panday, M., Asokan, R., Vijaya, L., Shantha, B., … der Steen, J. (2014). Comparison of saccadic reaction time between normal and glaucoma using an eye movement perimeter. *Indian Journal of Ophthalmology, 62*(1), 55. https://doi.org/10.4103/0301-4738.126182

McCulloch, C. E., & Neuhaus, J. M. (2011). Misspecifying the shape of a random effects distribution: Why getting it wrong may not matter. *Statistical Science, 26*(3). https://doi.org/10.1214/11-STS361

McTrusty, A. D., Cameron, L. A., Perperidis, A., Brash, H. M., Tatham, A. J., Agarwal, P. K., Murray, I. C., Fleck, B. W., & Minns, R. A. (2017). Comparison of Threshold Saccadic Vector Optokinetic Perimetry (SVOP) and Standard Automated Perimetry (SAP) in Glaucoma. Part II: Patterns of Visual Field Loss and Acceptability. *Translational Vision Science & Technology*, *6*(5), 4. https://doi.org/10.1167/tvst.6.5.4.

Mulligan, J., Stevenson, S., & Cormack, L. (2013). *Reflexive and voluntary control of smooth eye movements*. *8651*, 86510Z. https://doi.org/10.1117/12.2010333.

Munoz, D. P., Broughton, J. R., Goldring, J. E., & Armstrong, I. T. (1998). Age-related performance of human subjects on saccadic eye movement tasks. *Experimental Brain Research, 121*(4), 391–400. https://doi.org/10.1007/s002210050473

Murray, I. C., Fleck, B. W., Brash, H. M., MacRae, M. E., Tan, L. L., & Minns, R. A. (2009). Feasibility of saccadic vector optokinetic perimetry: A method of automated static perimetry for children using eye tracking. *Ophthalmology, 116*(10), 2017–2026.

Orban de Xivry, J.-J., Bennett, S. J., Lefèvre, P., & Barnes, G. R. (2006). Evidence for synergy between saccades and smooth pursuit during transient target disappearance. *Journal of Neurophysiology, 95*(1), 418–427. https://doi.org/10.1152/jn.00596.2005

Pel, J. J. M., van Beijsterveld, M. C. M., Thepass, G., & van der Steen, J. (2013). Validity and repeatability of saccadic response times across the visual field in eye movement perimetry. *Translational Vision Science & Technology, 2*(7), 3. https://doi.org/10.1167/tvst.2.7.3

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Redmond, T., Garway-Heath, D. F., Zlatkova, M. B., & Anderson, R. S. (2010). Sensitivity loss in early glaucoma can be mapped to an enlargement of the area of complete spatial summation. *Investigative Opthalmology & Visual Science, 51*(12), 6540. https://doi.org/10.1167/iovs.10-5718

Roth, H. L., Lora, A. N., & Heilman, K. M. (2002). Effects of monocular viewing and eye dominance on spatial attention. *Brain: A Journal of Neurology, 125*(Pt 9), 2023–2035. https://doi.org/10.1093/brain/awf210

Rountree, L., Mulholland, P. J., Anderson, R. S., Garway-Heath, D. F., Morgan, J. E., & Redmond, T. (2018). Optimising the glaucoma signal/noise ratio by mapping changes in spatial summation with area-modulated perimetric stimuli. *Scientific Reports, 8*(1), 2172. https://doi.org/10.1038/s41598-018-20480-4

Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., … Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution, 11*(9), 1141–1152. https://doi.org/10.1111/2041-210X.13434

Sharpe, J. A., & Sylvester, T. O. (1978). Effect of aging on horizontal smooth pursuit. *Investigative Ophthalmology & Visual Science, 17*(5), 465–468.

Soans, R. S., Renken, R. J., John, J., Bhongade, A., Raj, D., Saxena, R., … Cornelissen, F. W. (2021). Patients prefer a virtual reality approach over a similarly performing screen-based approach for continuous oculomotor-based screening of glaucomatous and neuro-ophthalmological visual field defects. *Frontiers in Neuroscience, 15*, Article 745355. https://doi.org/10.3389/fnins.2021.745355

Vonthein, R., Rauscher, S., Paetzold, J., Nowomiejska, K., Krapp, E., Hermann, A., … Schiefer, U. (2007). The normal age-corrected and reaction time-corrected isopter derived by semi-automated kinetic perimetry. *Ophthalmology, 114*(6), 1065–1072. e2. https://doi.org/10.1016/j.ophtha.2006.09.030

Zur, D., & Ullman, S. (2003). Filling-in of retinal scotomas. *Vision Research, 43*(9), 971–982. https://doi.org/10.1016/S0042-6989(03)00038-5