ORIGINAL PAPER

Free visual exploration of natural movies in schizophrenia

Johanna Elisa Silberg[1] · Ioannis Agtzidis[2] · Mikhail Startsev[2] · Teresa Fasshauer[1] · Karen Silling[1] ·
Andreas Sprenger[3] · Michael Dorr[2] · Rebekka Lencer[1,4]

Abstract

Background Eye tracking dysfunction (ETD) observed with standard pursuit stimuli represents a well-established biomarker for schizophrenia. How ETD may manifest during free visual exploration of real-life movies is unclear.
Methods Eye movements were recorded (EyeLink®1000) while 26 schizophrenia patients and 25 healthy age-matched controls freely explored nine uncut movies and nine pictures of real-life situations for 20 s each. Subsequently, participants were shown still shots of these scenes to decide whether they had explored them as movies or pictures. Participants were additionally assessed on standard eye-tracking tasks.
Results Patients made smaller saccades (movies ($p = 0.003$), pictures ($p = 0.002$)) and had a stronger central bias (movies and pictures ($p < 0.001$)) than controls. In movies, patients' exploration behavior was less driven by image-defined, bottom-up stimulus saliency than controls ($p < 0.05$). Proportions of pursuit tracking on movies differed between groups depending on the individual movie (group*movie $p = 0.011$, movie $p < 0.001$). Eye velocity on standard pursuit stimuli was reduced in patients ($p = 0.029$) but did not correlate with pursuit behavior on movies. Additionally, patients obtained lower rates of correctly identified still shots as movies or pictures ($p = 0.046$).
Conclusion Our results suggest a restricted centrally focused visual exploration behavior in patients not only on pictures, but also on movies of real-life scenes. While ETD observed in the laboratory cannot be directly transferred to natural viewing conditions, these alterations support a model of impairments in motion information processing in patients resulting in a reduced ability to perceive moving objects and less saliency driven exploration behavior presumably contributing to alterations in the perception of the natural environment.

Keywords Dynamic scenes · Visual exploration pattern · Smooth pursuit eye movements · Saliency · Schizophrenia · Visual processing

✉ Rebekka Lencer
rebekka.lencer@ukmuenster.de

1 Department of Psychiatry and Psychotherapy, University of Muenster, Albert-Schweitzer-Campus 1, Geb. A9, 48149 Muenster, Germany

2 Chair of Human-Machine Communication, Technical University Munich, Munich, Germany

3 Department of Neurology and Institute of Psychology II, University of Luebeck, Lübeck, Germany

4 Otto-Creutzfeldt Center for Cognitive and Behavioral Neuroscience, University of Muenster, Muenster, Germany

Introduction

Detailed observation of eye movements, with the eyes as a directly observable part of the brain, can provide specific information about brain function and dysfunction [25, 44]. In case of schizophrenia, numerous studies have shown disturbances of smooth pursuit eye movements (SPEM), so-called eye tracking dysfunction (ETD) [31]. More precisely, this deficit involves impairments of initiating and maintaining smooth pursuit of small moving objects with the eyes [10, 17, 18, 29, 45, 48]. One major factor contributing to ETD in patients is abnormal visual motion information processing and sensorimotor transformation as has been concluded from psychophysiological [9, 29, 48] and brain imaging studies showing altered activity in motion sensitive visual area V5 and posterior parietal association

cortex in patients associated with ETD [36–38]. Other studies extended these findings reporting similar ETD in unaffected first-degree relatives supporting a model of ETD as a biomarker for schizophrenia and psychotic disorders more generally indicating increased susceptibility to the disease [16, 30, 33].

To study the nature of ETD, most laboratory studies use abstract stimuli with single dots moving at well-defined velocities in the horizontal or vertical plane holding the advantage of reducing confounding factors to a minimum. However, in real-life situations, pursuit eye movements do not occur in isolation but are integrated in sequences of saccades and fixations depending on the subject's focus of attention. A consequently following question is whether and how ETD in patients may manifest in real-life situations?

To transfer visual exploration behavior from laboratory to more complex real-life situations, some studies have used well-defined real-life static pictures reporting that under free viewing circumstances schizophrenia patients perform smaller and fewer volitional saccades and fewer but longer fixations on objects of their interest: Typically, this scanning behavior results in restricted saccadic exploration patterns in patients [2, 3, 5, 32, 42, 53]. Other studies additionally showed patients differing from controls with respect to their spatial allocation of attention [42]. Two studies used a portable eye tracker to record patients' gaze behavior under head-free conditions. Delerue and co-workers studied fixation behavior while patients watched static objects on a table and conducted familiar (sandwich making) and unfamiliar (model building) tasks [11]. Dowiasch et al. [13] investigated visual tracking of stationary targets during self-motion induced smooth pursuit like eye movements as part of the slow phase of the vestibulo-ocular reflex (VOR), which was not altered in patients. They hypothesized that in real-life situations smooth eye movements are rather not altered due to the fact that patients can compensate a motion processing deficit by higher task engagement and by means of additional sensory information derived from vestibular signals or optic flow including the possibility of a freely movable head [13]. This hypothesis is supported by earlier fMRI studies reporting stronger activation in extraretinal brain systems in patients than controls during SPEM [36, 38].

But still, it remains unclear if and how ETD manifests during free viewing of real-life situations that involve moving targets. Therefore, the present study was designed to evaluate for the first time patients' free viewing behavior using uncut real-life movies, i.e., natural dynamic scenes that evoke both smooth pursuit and saccadic exploration behavior. For that, first, we determined the proportions of SPEM induced by moving objects in a scene. To identify smooth pursuit as a part of free viewing behavior, we used an algorithm recently developed by our group that has been shown to be much more sensitive than other state-of-the-art

procedures [1]. Second, besides SPEM we compared saccade and fixation metrics as well as spatial gaze distribution observed during free visual exploration between movies and static pictures in patients and controls to evaluate the motion effect on visual exploration behavior. Third, we examined to what extent visual exploration behavior was driven by image-defined, bottom-up saliency features of the scenes to study the question whether patients show alterations of bottom-up sensory information processing, especially when exploring dynamic scenes.

Methods

Twenty-six patients with a schizophrenia spectrum disorder according to DSM V ($N = 16$ schizophrenia, $N = 10$ schizoaffective disorder) participated in this study. Matching the patients with age, sex and education we recruited 25 healthy controls by newspaper advertisement and postings on noticeboards. Diagnoses were established by consensus processes using all available clinical information. Patients were clinically stable and receiving constant psychopharmacological treatment for at least two weeks with corresponding olanzapine equivalents [19] not exceeding 50 mg [Amisulpride ($n = 5$), Aripiprazole ($n = 12$), Clozapine ($n = 2$), Haloperidol ($n = 2$), Perazine ($n = 1$), Perphenazine ($n = 1$), Olanzapine ($n = 3$), Quetiapine ($n = 12$), Risperindone ($n = 5$), Ziprasidone ($n = 1$)]. Some patients additionally received antidepressants ($n = 16$) or mood stabilizers ($n = 5$). For more details see Supplementary Table 1.

Inclusion criteria for all participants included (a) age 18–55 years, (b) no history of serious head injury or systemic disease with effects on the central nervous system, (c) normal or corrected to normal vision (0.8–1.0 assessed by Landolt-C rings) and no evidence for red–green-deficiency [4], (d) no substance (cannabis, alcohol, synthetic drugs) dependency or abuse during the last 6 months. Controls were only included if the structured clinical interview for DSM IV (SCID) [54] was without evidence for any lifetime axis I disorder and if they denied any first-degree relatives with a diagnosis of a psychiatric disease. Current symptoms and cognitive status were assessed using standard rating scales, Table 1.

The ethics committee of the University Muenster approved the study. Both patients and controls gave their written informed consent before participating.

Eye movement recording

All eye movement recordings were conducted in the same room at the University Hospital in Muenster under constant conditions regarding illumination (0.01 cd/m^2), noise and standardized task instructions. Participants were seated at

Table 1 Demographic and clinical characteristics

Categories	Patients (N = 26)	Controls (N = 25)	p (T-test)
Age, mean (SD)	38.3 (10.8)	37.0 (9.2)	0.65
Sex (female|male)	10|16	12|13	0.50
Visual acuity, mean (SD)	1.01 (0.1)	0.96 (0.1)	0.11
Verbal IQ (MWT-B raw score) [28], mean (SD)	28.7 (4.8)	30.6 (4.4)	0.15
Duration of illness (years)[a], mean (SD)	10.8 (7.9)	n.a	n.a
Montgomery Asberg Depression Rating Scale [34], (MADRS), mean (SD)	12.65 (6.7)	0.84 (1.8)	< 0.001
Global Assessment of Functioning (GAF) [14], mean (SD)	56.27 (13.8)	n.a	n.a
Positive and Negative Syndrome Scale (PANSS) [27]			
Positive, mean (SD)	12.50 (4.9)	n.a	n.a
Negative, mean (SD)	14.54 (6.4)	n.a	n.a
General, mean (SD)	31.38 (9.0)	n.a	n.a
Global, mean (SD)	58.73 (18.8)	n.a	n.a
Olanzapine equivalents (mg) [19], mean (SD)	21.42 (12.7)	n.a	n.a

n.a not applicable

[a]Since diagnosis was first established by a psychiatrist or otherwise experienced medical doctor

75 cm distance from a CRT monitor (Samsung SyncMaster 950p Plus, diagonal of 45.5 cm, resolution of 1280 × 1024 pixels) with their heads stabilized by a chin rest with forehead restraint. Stimuli were presented at a 1280 × 720 pixel resolution, corresponding to 27.9° × 15.7° visual angle. For eye movement recordings the desktop mount *EyeLink®1000* system (SR Research Ltd., Mississauga, Ontario, Canada) was used. Data was recorded at 500 Hz binocularly and recordings of the eye providing the best data quality were used for analysis. This was checked systematically for all recordings. The experiment itself consisted of three tasks that followed in the order of task 3, 1, 2. Stimuli were generated by custom-made programs using C++ (ISO Standard ISO/IEC 14882:2003) for task 1 and Psychtoolbox (Version 3.10, Brainard 1997) for *Matlab®* (R2010a, The Mathworks INC., Natic, MA, USA) [7] for task 3.

Task 1: visual exploration of dynamic and static real-life situations

Stimulus material were movies of real-life outdoor situations from the GazeCom dataset (http://www.inb.uni-luebeck.de/tools-demos/gaze) [12] that consists of 18 movies lasting 20 s each. We selected nine of these as movies for dynamic exploration trials (breite_strasse, bridge_2, doves, ducks_boat, golf, koenigstrasse, puppies, st_petri_mcdonalds, street). The remaining nine movies were presented as static pictures, i.e., representative still shots (beach, bridge_1, bumblebee, ducks_children, holsten_gate, roundabout, sea, st_petri_gate, st_petri_market) for 20 s each, Fig. 1. The allocation of GazeCom movies to either dynamic or static condition was performed manually, aiming to use those

movies that had less motion to begin with for the static condition. No sound was played. Participants were instructed to watch attentively imagining they were direct observers of the scenes. They were informed to be asked questions about its content in task 2. Stimuli were presented in a pseudo-randomized order. There was a five-point eye tracker calibration preceding task presentation. During the task, each trial was preceded by a one-point drift correction.

Task 2: recall of motion information

Representative still shots of movies and pictures from task 1 were presented for five seconds each by a Powerpoint® presentation in pseudo-randomized order different from task 1. Participants had to decide whether the still shots had been part of a movie or picture. Eye movements were not recorded during task 2.

Task 3: standard saccade and pursuit stimuli

Participants were tested on standard stimuli including (a) visually guided reflexive saccades of 4°, 8° and 12° to left and right in randomized order (18 trials), and (b) smooth pursuit of triangular wave tasks. The latter consisted of blocks of constant velocity ramps of either 8°/s or 16°/s each block containing eight target sweeps from left to right and vice versa in the horizontal plane. Both blocks were presented separately. There was a nine-point eye tracker calibration preceding task 3. Each block was preceded by a one-point drift correction.

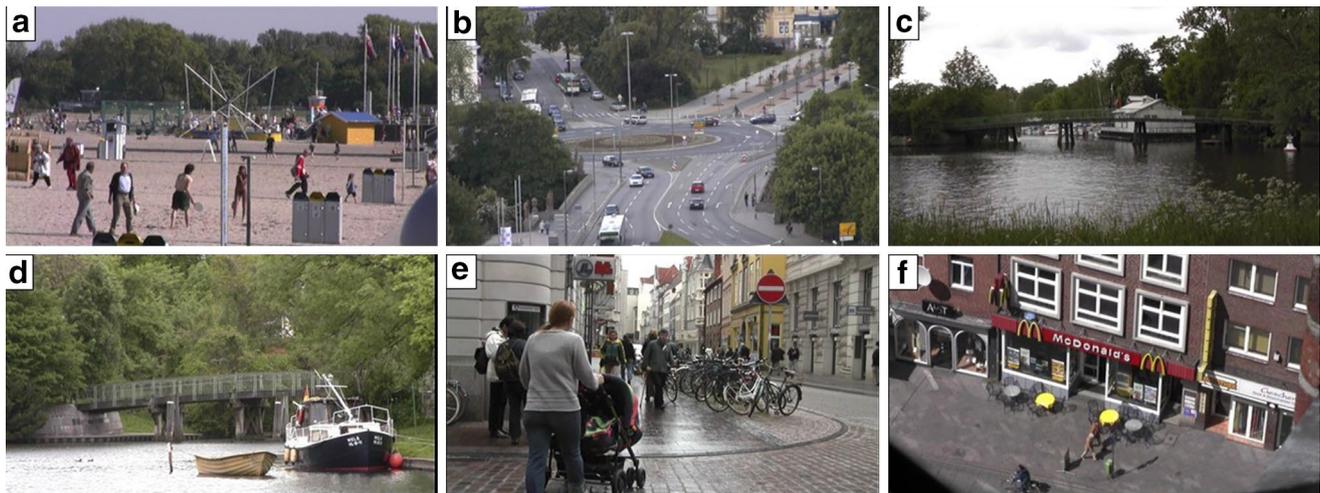The complete eye testing procedures took about 25 min including eye tracker calibration.

Fig. 1 Examples of stimuli shown in task 1. First row— shows pictures: **a** beach, **b** roundabout, **c** bridge_1; Second row— shows movie still shots: **d** ducks_boat, **e** koenigstrasse, **f** st_petri_mcdonalds

Data processing

An overview over measures used in this study is available in Supplementary Table 2.

Task 1: After data pre-processing, analyses of task 1 included the following steps: (1) determine SPEM proportions of visual exploration behavior, (2) determine spatial gaze distribution to identify possible central bias effects (3) determine the similarity of spatio-temporal gaze distributions between observers, and (4) determine to what extent exploration behaviour was driven by image-defined, bottom-up stimulus saliency features.

Details of data pre-processing have been described previously including detection and removal of invalid samples and blink artifacts from recordings [12]. Saccades were identified by a dual-threshold approach to distinguish saccades from noise, and fixations were identified based on the dispersion of inter-saccadic samples. After that, smooth pursuit episodes were detected using the recently developed "multi-observer smooth pursuit" (MOSP) method [1] (a toolbox is publicly available at http://michaeldorr.de/smoothpursuit). MOSP works as follows: because of the relatively slow speed, SPEMs are harder to distinguish from noisy fixations than saccades, and this is particularly true when using natural instead of synthetic stimuli, where the target-trajectory ground truth is not known a priori. MOSP operates on the assumption that SPEM must be induced by image motion, and should, therefore, occur in similar spatiotemporal locations for different observers, whereas noise should be independent across observers. By integrating information across multiple observers, MOSP is more robust against noise and thus achieves higher precision and recall on a hand-labelled "ground truth" dataset than state-of-the-art algorithms [43].

In detail, at first, gaze data from all participants were pooled for each stimulus, and then all episodes that could be reliably labeled as either saccade or fixation were removed. A spatio-temporal clustering was performed on the remaining samples to identify episodes of highly similar gaze traces (that were neither fixations nor saccades). The criterion for cluster identification was based on the DBSCAN [15] algorithm, with a spatial radius of 2° of visual field; the free parameter minPts was set so that either one observer would have to maintain a perfect (i.e., continuous) SPEM-like eye movement for at least 160 ms, or, which is more likely, at least two observers' gaze traces would have to be part of a cluster. All the gaze samples inside the cluster were labelled as SPEM.

Second, we quantified how much time observers spent on which location while exploring the stimuli. "Central bias" explains the tendency of observers to preferentially fixate near the center of a scene [46]. Central bias was quantified by calculating the Euclidean distance to the mean gaze position per observer and stimulus for each gaze sample. The histogram over these distances was computed with bins of 10 pixels width (approx. 0.2°). To account for the central starting position imposed by the drift correction before each trial, we discarded the first 500 ms of gaze data for this analysis.

Third, we measured the similarity of spatio-temporal gaze distributions between observers to quantify how homogeneous scan paths were within groups and between different stimulus types. This measure indicates to what extent the order of fixated spatial locations is similar between subjects. For that we computed the normalized scanpath saliency (NSS) as follows [12]: a fixation map was created from spatio-temporally smoothed gaze data of all but one observer for a temporal window corresponding to one movie frame

(33 ms), followed by normalization to zero mean, unit standard deviation. The values of this fixation map at the locations of gaze of the left-out observer were then averaged. This process was repeated over all movie frames and observers. Positive NSS scores indicate similarity, whereas a score of 0 indicates no systematic relationship between observers' gaze patterns.

Fourth, we related the physical stimulus saliency to eye movement predictability to evaluate to what extent observers' viewing behavior was driven by bottom-up information processing. For that, two image-based measures of "saliency'' were computed that aim at quantifying the informativeness of an image location, e.g., a moving object tends to attract human attention, because it is more informative than a uniform background. A high predictability of eye movements based on such measures indicates stronger bottom-up driven attentional control (based on image features such as contrasts and motion), whereas lower predictability indicates a stronger contribution of top-down, volitional mechanisms of attention. Specifically, for each gaze location we calculated the "Ensemble of Deep Networks" (eDN) measure [51] (on static pictures) or the geometrical invariant K of the structure tensor [52] (on movies); both have been shown to perform well on standard computer vision benchmarks [8, 52].

Task 2: Parameter of interest was the rate of correctly identified still shots as movies or pictures that had been taken from stimuli used in task 1, in the following referred to as recall score.

Task 3: Eye movement recordings were calibrated, saccades and blinks were automatically detected and manually checked with an interactive program written in Matlab ® R2010a (The MathWorks, Natick, MA, USA). In saccade tasks, parameters of interest were the latency of the first saccade (with a minimal latency of 130 ms), and its amplitude to calculate gain (saccade amplitude/target amplitude).

In pursuit tasks, eye velocity was calculated as the first derivative of the eye position signal (filtered with a Gaussian low-pass filter with − 3 db at 100 Hz) after blinks and saccades had been removed. Parameter of interest was maintenance gain (mean eye velocity/mean target velocity) in an interval of 450–2400 ms (8°/s) or 225–1200 ms (16°/s) in the middle of each ramp. In each participant, grand medians over all trials were determined in 8°/s and 16°/s blocks, respectively.

Statistical analyses

Statistical analyses were carried out in SPSS® (version 24, IBM Deutschland GmbH, Ehningen, Germany) and the R® statistical package. We used mixed model repeated measures analysis of variances (ANOVA) for data collected in task 1 and 3.

Task 1: Picture and movie trials were analyzed separately to study effects on fixation rate and duration, saccade rate and amplitude, and proportion of SPEM. For that, individual measures for each trial and subject were entered in ANOVAs. "Trial" was used as within-subject factor, while "group" represented the between-subject factor.

To assess group differences on central bias, we calculated the bin-wise histogram difference between patients and controls for pictures and movies, respectively. We also calculated this difference in central bias between the two stimulus types, i.e., movies and pictures, Fig. 3.

To compare spatiotemporal gaze distributions between groups, NSS scores were pooled over time and over pictures and movies, respectively. Group differences were assessed by Kolmogorov–Smirnov test, Fig. 4.

To test for differences in saliency-related visual exploration behavior, saliency values were binned by their distance to the stimulus center. For each of the 101 bins, the mean saliency value was computed, and finally, the bin means were fitted with a LOESS function, using R default parameters. A bootstrapping procedure was performed to assess statistical significance. For 2000 iterations, subsets of 5000 gaze samples each from patients and controls and their corresponding saliency values were drawn and fitted with a LOESS function over distance to the stimulus center. Eccentricities, where more than 97.5% (less than 2.5%) of the iterations showed a higher (lower) fitted saliency value were deemed to differ significantly, Fig. 5.

Task 2: Recall scores were compared between groups by t-tests.

Task 3: Target amplitude (4°, 8°, 12°) in standard saccade tasks and target velocity (slow: 8°/s, fast: 16°/s) in standard pursuit tasks were defined as within-subject factors, respectively. "Group" was defined as between-subject factor to test for differences in saccade latency, saccade amplitude gain and pursuit velocity gain between groups.

Pearson's correlations were used first, to relate eye movement measures from task 1 and 3 to clinical variables as listed in Table 1, e.g., PANSS scores, and second, to evaluate possible influences of antipsychotic medication on eye movement measures. The latter showed that the following eye movement parameters correlated with individual olanzapine equivalents in patients: saccade latency in task 3 ($r = -0.41$, $p = 0.04$) and fixation duration (static: $r = 0.58$, $p < 0.01$, dynamic: $r = 0.50$, $p < 0.01$) and fixation rates in task 1 (static $r = -0.48$, $p = 0.01$, dynamic: $r = -0.49$, $p = 0.01$). Accordingly, these saccade and fixation parameters were corrected using the residuals from a regression approach before entering them in ANOVA. We will report Greenhouse Geisser corrected p values and effect sizes, but uncorrected degrees of freedom.
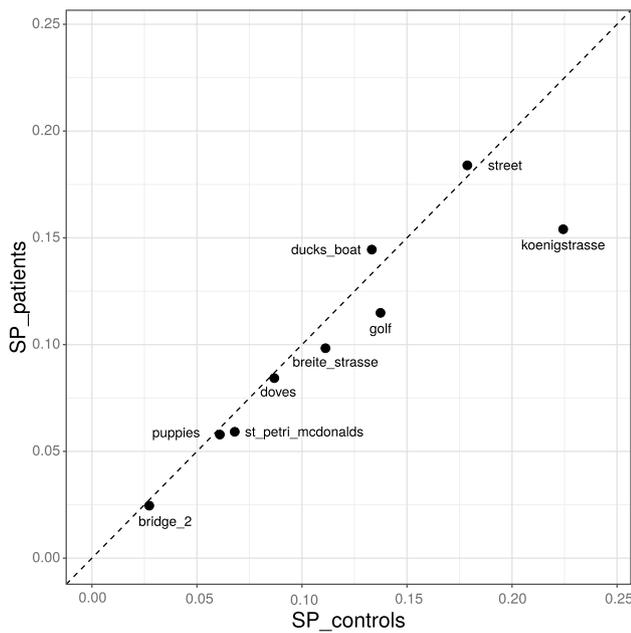
Fig. 2 Proportions of smooth pursuit eye movements (SP) in patients ($N = 26$) and healthy controls ($N = 25$) while exploring movies of natural scenes from GazeCom (20 s). Each dot represents one of the nine movies. Koenigstrasse was the movie in which patients performed considerably less SPEM than healthy controls

Results

Task 1: exploration behavior of natural movies and pictures

Exploration behavior during movies

Mean durations of the automatically detected smooth pursuit episodes were between 150 and 200 ms. Such short time intervals rather correspond to the pursuit initiation phase but hardly represent pursuit maintenance. Proportions of pursuit during visual exploration differed clearly between trials, i.e. individual movies (movie: $F_{(8,392)} = 64.52$, $p < 0.001$) and groups depending on the individual movie (group × movie $F_{(8,392)} = 3.09$, $p = 0.011$), Fig. 2. The largest difference in pursuit proportions between groups was observed with movie koenigstrasse. Follow-up correlation analyses showed that pursuit proportions during exploration of movies were not related to pursuit maintenance gain from standard pursuit stimuli.

Fixation rates and durations, as well as saccade rates and amplitudes also differed between the individual movies (for all parameters $p < 0.001$). Saccade amplitudes were smaller in patients than in controls ($F_{(1,49)} = 9.46$, $p = 0.003$), Table 2, whereas saccade and fixation rates and fixation durations neither differed between groups nor were there any interactions of group x movie.

Table 2 Eye movement parameters on movies and pictures (task 1) and basic standard stimuli (task 3) and results of recall motion task (task 2), means with standard deviation (SD)

	Patients	Controls
Task 1: eye movement parameter on movies		
Fixation rate (per sec.)	1.99 (0.38)	2.08 (0.24)
Fixation duration (ms)	461.02 (113.16)	419.19 (45.35)
Saccade rate (per sec.)	1.64 (0.38)	1.67 (0.29)
Saccade amplitude (°)	4.28 (0.73)	4.84 (0.56)
Proportion of smooth pursuit (%)	10.2 (2.9)	11.4 (1.9)
Task 1: eye movement parameter on pictures		
Fixation rate (per sec.)	2.33 (0.36)	2.41 (0.31)
Fixation duration (ms)	397.79 (90.54)	375.53 (52.08)
Saccade rate (per sec.)	1.86 (0.41)	1.91 (0.33)
Saccade amplitude (°)	3.92 (0.67)	4.50 (0.56)
Task 2: recall score		
Pictures	0.91 (0.12)	0.96 (0.11)
Movies	0.88 (0.15)	0.95 (0.10)
Total	0.90 (0.10)	0.95 (0.08)
Task 3: standard stimuli		
Saccade latency (ms)	168.08 (21.6)	177.4 (24.3)
Saccade gain	0.98 (0.10)	0.99 (0.08)
Smooth pursuit gain	0.90 (0.07)	0.94 (0.05)

Exploration behavior on pictures

As for movies, saccade rates and amplitudes, as well as fixation rates and duration differed between trials, i.e. the individual pictures (all $p < 0.001$). Again, patients made smaller saccades than controls ($F_{(1,49)} = 10.97$, $p = 0.002$), Table 2. Group differences did not differ between the individual pictures (group × picture $p > 0.05$).

Comparison of eye movement parameters between movies and pictures showed that in movies fixation and saccade rates were lower, fixation durations were longer and saccade amplitudes were larger than during exploration of pictures (all $p < 0.001$), Table 2. However, this clear effect of motion information on visual exploration behavior did not differ between groups, i.e., there were no interactions of group x trial type for any eye movement parameter.

Spatial gaze distribution

On both pictures and movies, patients showed a stronger central bias than controls ($p < 0.001$) meaning their gaze was more often directed towards the center of a scene, i.e., at eccentricities $< 5°$, Fig. 3. Comparing gaze guidance between movies and pictures across groups (blue graph) showed that dynamic scenes attenuated the central bias effect seen with static pictures so that more time was spent between $5°$ and $10°$ eccentricity.

Similarity of spatio-temporal gaze distribution

As expected, between-subject homogeneity as indicated by NSS scores was higher on movies than pictures ($p < 0.001$) reflecting that temporal information, i.e., changing image regions in dynamic scenes, has a strong effect on gaze guidance, Fig. 4. However, patients' viewing behavior on pictures was significantly more similar than controls', i.e., NSS scores were significantly higher than those in controls ($p < 0.001$). It turned out to be exactly the opposite for movies where controls showed significantly higher NSS scores than patients ($p < 0.001$), Fig. 4.

Stimulus saliency driving exploration behavior

Concerning the question whether exploration behavior was driven by bottom-up guidance or rather top-down control, our findings point to patients being less influenced by bottom-up saliency on movies than controls at eccentricities of less than $5°$ ($p < 0.05$). This group difference was not observed for pictures, Fig. 5.
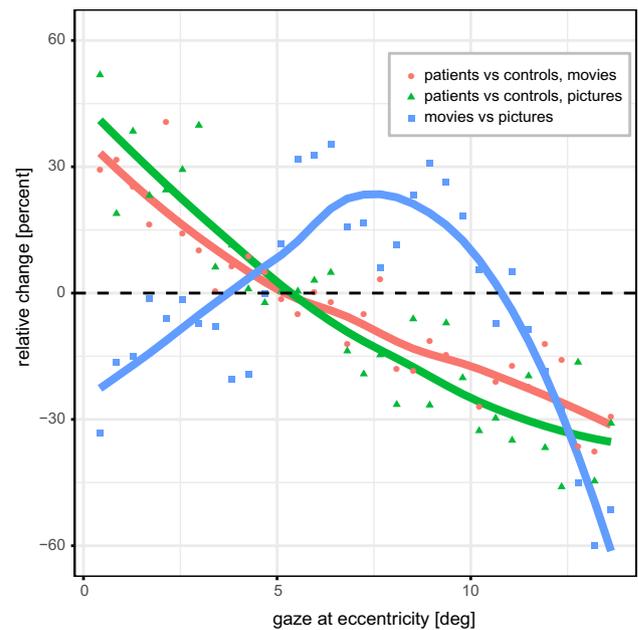

Fig. 3 Comparison of the central bias in patients ($N = 26$) and healthy controls ($N = 25$) for natural movies and pictures. Each dot represents the ratio of time spent at a certain eccentricity with an eccentricity of 0 indicating the center of a scene. Red and green dots above the dotted line (relative change > 0) indicate that patients spent more time at a certain eccentricity. This was the case for eccentricities $< 5°$ reflecting a stronger central bias in patients. Blue dots above the line indicate that more time was spent at a certain eccentricity when exploring a movie. Solid lines indicate LOESS function fits

Task 2: recall of motion information

Rates of still shots correctly identified as movies or pictures differed significantly between groups ($T_{(49)} = -2.054$, $p = 0.046$), Table 2. Neither performance on standard pursuit stimuli nor proportions of smooth pursuit while exploring a movie correlated with movie recall scores. For more details see Supplementary Table 3.

Task 3: standard stimuli

Reflexive saccades: Across trials, amplitude gain and latency did not differ between groups ($p > 0.05$), Table 2.

Smooth pursuit gain: Across participants, eye velocity gain decreased with increasing target velocity ($F_{(1,48)} = 4.48$, $p = 0.039$), Table 2. As expected, patients had lower SPEM gains compared to controls ($F_{(1,48)} = 5.05$, $p = 0.029$), Table 2. There was no interaction of group by target speed.

Relation of eye movement parameters and clinical ratings

There were only two clinical variables that were related to eye movement measures. In patients, smaller proportions
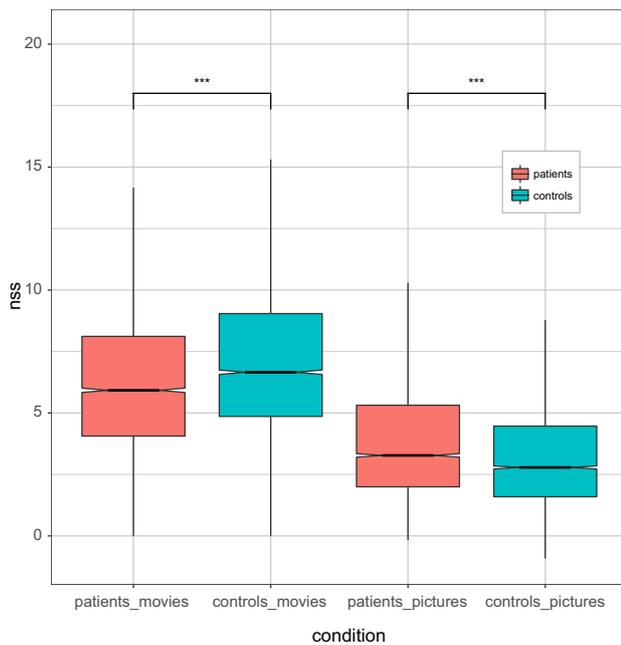
Fig. 4 Variability of scan paths of patients ($N=26$, red boxplots) and controls ($N=25$, blue boxplots) on movies and pictures. NSS score is plotted on the ordinate as absolute values, 0 is indicating totally different scan paths, while higher values indicate growing homogeneity of scan paths within groups (patients and controls on movies and pictures). Boxplots show that variability of the individual scan paths in patients is significantly higher on movies and lower on pictures than scan paths homogeneity in controls, *** $p < 0.001$

of pursuit were correlated with longer duration of illness ($r = -0.45$, $p = 0.021$) and higher PANSS negative scores were correlated with lower saccade rates on movies ($r = -0.42$, $p = 0.023$) but not on pictures. No other clinical characteristics as listed in Table 1 were related to any other eye movement parameter.

Discussion

ETD is a repeatedly replicated neurophysiological deficit observed in schizophrenia [31]. Extending previous eye movement studies with schizophrenia patients, we here investigated whether and how ETD may manifest during free visual exploration of uncut real-life movies. Generally, when analyzing free viewing behavior on movies of real-life situations, it has to be considered that first, motion increases complexity of stimuli but in turn strongly attracts the observers' attention and gaze, because it usually marks the spot holding the most relevant information [41]. Second, movies as provided by the GazeCom dataset were taken from familiar surroundings, meaning that moving objects occurring in the movies, i.e., a person on a bicycle, may have drawn less attention than synthetic stimuli used in laboratories when subjects are explicitly instructed to hold attention on a certain target and follow its movement [21].

Our main findings show that SPEM proportions while watching a movie were generally low (patients: 10.2% (SD:
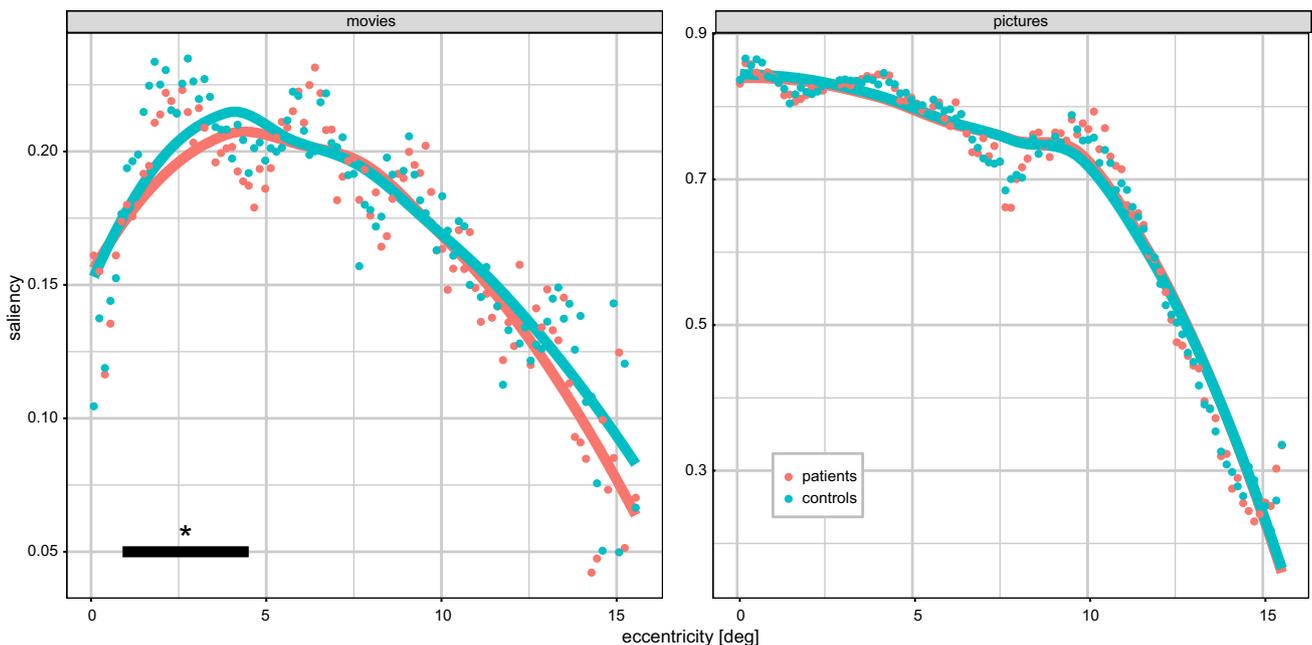

Fig. 5 Mean saliency values summed up over all saliency maps of the nine movies (left figure) and nine pictures (right figure), separately for patients ($N=26$, red graphs) and healthy controls ($N=25$, blue graphs). These graphs show that for both patients and controls, attended locations in the periphery were less salient than central locations. In movies, patients' central fixations ($<5°$) were less driven by saliency than those of controls (Note the differently scaled ordinates for movies and pictures)

2.8%), controls: 11.4% (SD: 1.9%)) and were even reduced in patients depending on the individual movie. Patients' visual exploration behavior was more restricted than in controls characterized by smaller saccades and a stronger central bias both on pictures and movies. Besides this, patients' gaze control when watching a movie was less strongly related to stimulus saliency than in controls suggesting that in patients volitional visual exploration behavior of movies is less driven by bottom-up motion information processing. In the following, we will discuss the implications of our findings for a better understanding of disturbances in sensorimotor systems in patients with schizophrenia.

The role of smooth pursuit while watching movies

Overall, pursuit episodes observed with movies were very short reflecting rather the pursuit initiation phase than pursuit maintenance, which is only established after 300 ms of SPEM. This observation supports a model of pursuit initiation being of higher importance during real-life situations than longer lasting phases of pursuit maintenance. According to Thorpe et al. [47] visual recognition happens within the first 150 ms, [47], which is in line with our finding of short pursuit episodes on movies. After this initiation interval pursuit is only maintained if there is a specific purpose in pursuing a moving target, e.g., when driving a vehicle [21, 22, 47]. Note that standard triangular wave tasks rather measure pursuit maintenance instead of pursuit initiation ability, which may explain why we could not find direct correlations between pursuit maintenance gain in task 3 and episodes of pursuit initiation during movies in task 1 in neither patients nor controls.

SPEM proportions highly differed between movies, Fig. 2. Higher predictability of familiarly moving features in real-life scenes may result in less SPEM proportions such as in movie bridge_2. Patients differed from controls in their SPEM proportions depending on the individual movie. This effect was more pronounced for movies evoking larger SPEM proportions indicating lower predictability, and most so in movie koenigstrasse. Note that longer duration of illness but not age was associated with smaller SPEM proportions suggesting a certain illness progression effect that may impact occipito-parietal networks shown previously to be altered during SPEM in patients with schizophrenia [36–38].

Higher predictability of the movie content could also facilitate acquired top-down driven capabilities and learned exploration strategies in patients to compensate for possible ETD [21, 57]. This hypothesis is in line with observations from studies using portable eye trackers in the natural environment that did not find alterations of SPEM like eye movements evoked by head movements in patients [13]. The impact of extraretinal mechanisms on SPEM generation has also been shown by laboratory SPEM studies that used highly predictable synthetic stimuli that were designed in a way that abrupt target reversals were omitted, so-called oscillating tasks. Using these tasks, ETD in patients was absent even at high target speed underlining the importance of top-down predictive mechanisms for SPEM control [29, 37].

Spatial gaze distributions while exploring movies and pictures

As expected, fixation and saccade rates, as well as fixation durations were lower on movies than on pictures while saccade amplitudes increased and gaze was directed more often to eccentric positions in movies than in pictures reflecting the strong effect of motion on gaze behavior (blue dots in Fig. 3). Our findings of a more restricted exploration pattern in patients characterized by smaller saccade amplitudes and a stronger central bias are in line with current research on free visual exploration of natural pictures [3, 32, 39, 40, 42]. Most importantly, we found a similar visual exploration impairment also for movies which was specifically correlated with higher negative symptom expression on the PANSS. Thus, although motion stimuli may be able to enhance attentional gaze to more eccentric locations even in patients, the motion effect does not seem strong enough to overcome generally restricted visual exploration patterns in patients which may be at least partly related to attenuated motivation and a diminished expression of psychological functions as reflected by the PANSS negative score.

The findings on spatial gaze distribution are complemented by our observations on between-subject scan path homogeneity (NSS scores) indicating whether subjects directed their gaze to specific locations of a scene in a similar order in time. Generally, between-subject homogeneity is known to be higher on movies than pictures, as we found in both patients and controls representing the known effect of motion on equalizing scan paths among observers [41]. Higher between-subject homogeneity on pictures in patients than controls presumably reflects their restricted exploration behavior. The opposite finding that the otherwise homogeneity enhancing effect of movies was weaker in patients than controls, i.e., NSS scores in patients were lower than in controls, underlines the hypothesis of alterations in brain systems for visual motion and sensorimotor processing affecting volitional visual exploration in patients. This hypothesis is in line with findings from functional imaging studies showing disturbances in motion-sensitive visual area V5 and posterior parietal association cortex in patients with schizophrenia [36–38]. Further support for impaired visual motion information processing in patients comes from our observation of reduced motion recall ability in patients revealed in task 2.

The influence of stimulus saliency on visual exploration behavior

The guided search model suggests [55] that under natural circumstances, saliency induces an initial draw of gaze and attention [24], followed by early recognition of potential attentional targets resulting in an adapted (top-down) gaze control, i.e., a selection of an eye movement/attentional target grounded on experience [24, 55, 57]. Defining saliency by the physical characteristics of stimuli we found that across participants, saliency effects were highest around the center of a scene for eccentricities $< 5°$ in movies and for eccentricities $< 10°$ in pictures, Fig. 5. Beyond these eccentricities, saliency effects generally decreased. While there was no difference in saliency effects on gaze behavior between groups for pictures, the effect of stimulus saliency in movies was reduced in patients for eccentricities between $0.5°$ and $4.5°$, thus in the region where stimulus saliency driving gaze behavior was strongest in controls. Relating these findings to the findings on spatial gaze distribution, we suggest that although patients spent more time at locations around the center of a scene at eccentricities $< 5°$, Fig. 3, this behavior does not seem to be related to saliency features. Together, these findings support the notion of alterations in bottom-up visual motion information processing resulting in enhanced top-down gaze control in patients.

Reduced stimulus saliency-related gaze behavior during movies in patients gives rise to the question whether our findings may reflect "aberrant saliency" mechanisms as defined in the model by Kapur [26]. This model suggests that due to imbalances in dopaminergic transmitter systems patients direct their attention to aspects in their environment that are generally of minor importance, i.e., are less salient, but get over-interpreted by patients and drive the development of positive symptoms. From our findings, it may be concluded that patients attribute aberrant saliency to objects around the center of a scene that healthy subjects rather consider as uninteresting. Patients apparently spend more time on processing what they see, so that even strong saliency features such as motion are not able to lead their gaze to more peripheral locations. Neurophysiological sensitivity to visual selective attention under high processing load has been suggested as a key feature predisposing to psychotic disorders [50]. Thus, sensory overstimulation by saliency features limiting the ability to focus and switch attention [6, 56] might as well contribute to a restricted volitional visual exploration behavior resulting in impairments of adapting eye movements to attentional demands [22].

Clinical implications

It should be considered that restricted visual exploration behavior and more top-down controlled viewing strategies in patients influence each other in both ways. Patients seem to have problems to release their attention from features in the center of a scene resulting in less scanning and reduced recognition of potential features of interest in the periphery. Transferring this to real-life visual exploration and perception, this may explain why patients have problems in processing contextual information, as proposed for social or emotional contents [20], i.e., resulting in a more local exploration pattern [23, 42, 49]. Consequently, they possibly miss information and in turn randomly overestimate and misinterpret distinct details of a scene without considering the context, i.e., patients may jump to conclusions before considering all aspects and information that is contained in a scene supporting the development of paranoia and other delusions [26]. These aspects of restricted visual exploration behavior that we found to be related to illness duration and negative symptom expression to some extent, should be considered by treatment approaches that aim at modifying maladaptive thoughts or perceptions. The metacognitive training (MCT) developed by Moritz and colleagues represents one treatment approach following these implications from neurophysiological studies [35]. For example, the MCT module "jumping to conclusions" focusses on systematic visual information gathering before drawing a definite conclusion about what is depicted in a scene.

Limitations

As usual, there are limitations to this studies that should be considered when interpreting the results. First, all patients were on stable medication excluding sedatives. We considered possibly relevant medication effects using a regression approach but cannot fully exclude medication as a confounding factor on our results as it is the case in all studies on medicated patients. Follow-up studies with unmedicated first-episode patients are required to (a) assess the effects of medication on free viewing behavior on real-life movies and (b) to get a more detailed idea of how the disease and medication possibly influence patients' perception of their natural surroundings. Second, our sample consisted of chronically ill patients with mild symptom expression, i.e., relatively low PANSS scores at the time of testing. This said, our findings rather reflect trait than state-dependent neurophysiological dysfunctions. Future studies should include patients with higher symptom expression to relate visual exploration deficits more specifically to positive symptoms such as delusions and hallucinations. Third, we only used a very simple motion recall task, which may explain why we did not find a more specific relation between recall ability and visual exploration behavior. Future studies should include more detailed and elaborate questions about the movie content and also a larger set of naturalistic stimuli. Information about

the participant's interpretation of the content of the movies and pictures, respectively, would allow for a more detailed explanation of the observed eye movement patterns. Additionally, eye movement recordings during the recognition process might provide further insight into motion memory and recognition. Fourth, a stronger central bias in patients and their more restricted exploration behavior may not be regarded independently but rather be interrelated. Follow-up studies examining this topic in more detail are needed. Fifth, a larger sample size is needed to identify differences between patient subgroups of the schizophrenia spectrum, i.e., schizophrenia and schizoaffective disorder.

Conclusions

Despite these limitations, this study provides novel information about visual exploration behavior of natural dynamic scenes in patients with schizophrenia. Although the well-established endophenotype of ETD observed under laboratory conditions cannot be directly transferred to natural viewing conditions, our findings support a model of impairments in motion information processing in patients that may result in a reduced ability to perceive moving objects and less saliency driven exploration behavior thus contributing to alterations in the perception of the natural environment.

References

1. Agtzidis I, Startsev M, Dorr M (2016) Smooth pursuit detection based on multiple observers. In: ACM Symposium on Eye Tracking Research and Applications ETRA '16: ACM Symposium on Eye Tracking Research and Applications ETRA '16. ACM, Charleston, SC, pp 303–306

2. Benson PJ, Beedie SA, Shephard E et al (2012) Simple viewing tests can detect eye movement abnormalities that distinguish schizophrenia cases from controls with exceptional accuracy. Biol Psychiatry 72(9):716–724. https://doi.org/10.1016/j.biopsych.2012.04.019

3. Bestelmeyer PEG, Tatler BW, Phillips LH et al (2006) Global visual scanning abnormalities in schizophrenia and bipolar disorder. Schizophr Res 87(1–3):212–222. https://doi.org/10.1016/j.schres.2006.06.015

4. Birch J (1997) Efficiency of the Ishihara test for identifying red-green colour deficiency. Ophthalmic Physiol Opt 17(5):403–408. https://doi.org/10.1016/S0275-5408(97)00022-7

5. Bradley MM, Houbova P, Miccoli L et al (2011) Scan patterns when viewing natural scenes: emotion, complexity, and repetition. Psychophysiology 48(11):1544–1553. https://doi.org/10.1111/j.1469-8986.2011.01223.x

6. Braff DL (1993) Information processing and attention dysfunctions in schizophrenia. Schizophr Bull 19(2):233–259

7. Brainard DH (1997) The psychophysics toolbox. Spat Vis 10(4):433–436. https://doi.org/10.1163/156856897X00357

8. Bylinskii Z, Judd T, Oliva A et al. What do different evaluation metrics tell us about saliency models? https://arxiv.org/abs/1604.03605

9. Chen Y-C, Lu Y-C, Lung F-W (2011) Confirmatory factor analysis in neurophysiological and neuropsychological dimensions of schizophrenia. Int J Neurosci 121(9):528–535. https://doi.org/10.3109/00207454.2011.582239

10. Clementz BA, McDowell JE (1994) Smooth pursuit in schizophrenia. Abnormalities of open- and closed-loop responses. Psychophysiology 31(1):79–86. https://doi.org/10.1111/j.1469-8986.1994.tb01027.x

11. Delerue C, Boucart M (2013) Visual exploration and action processing in schizophrenia. Cogn Neuropsychiatry 18(3):153–167. https://doi.org/10.1080/13546805.2012.683245

12. Dorr M, Martinetz T, Gegenfurtner KR et al (2010) Variability of eye movements when viewing dynamic natural scenes. J Vis 10(10):28. https://doi.org/10.1167/10.10.28

13. Dowiasch S, Backasch B, Einhäuser W et al (2016) Eye movements of patients with schizophrenia in a natural environment. Eur Arch Psychiatry Clin Neurosci 266(1):43–54. https://doi.org/10.1007/s00406-014-0567-8

14. Endicott J, Spitzer RL, Fleiss JL et al (1976) The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. Arch Gen Psychiatry 33(6):766–771

15. Ester M, Kriegel H-P, Sander J et al (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the second internationalconference on knowledge discovery and data mining KDD'96. AAAI Press, pp 226–231

16. Ettinger U, Kumari V, Crawford TJ et al (2004) Smooth pursuit and antisaccade eye movements in siblings discordant for schizophrenia. J Psychiatr Res 38(2):177–184

17. Farber RH, Clementz BA, Swerdlow NR (1997) Characteristics of open- and closed-loop smooth pursuit responses among obsessive-compulsive disorder, schizophrenia, and nonpsychiatric individuals. Psychophysiology 34(2):157–162. https://doi.org/10.1111/j.1469-8986.1997.tb02126.x

18. Friedman L, Jesberger JA, Siever LJ et al (1995) Smooth pursuit performance in patients with affective disorders or schizophrenia and normal controls: analysis with specific oculomotor measures, RMS error and qualitative ratings. Psychol Med 25(2):387–403

19. Gardner DM, Murphy AL, O'Donnell H et al (2010) International consensus study of antipsychotic dosing. Am J Psychiatry 167(6):686–693. https://doi.org/10.1176/appi.ajp.2009.09060802

20. Green M, Uhlhaas P, Coltheart M (2005) Context processing and social cognition in schizophrenia. CPSR 1(1): 11–22. https://doi.org/10.2174/1573400052953529

21. Hayhoe M, Ballard D (2005) Eye movements in natural behavior. Trends Cogn Sci 9(4):188–194. https://doi.org/10.1016/j.tics.2005.02.009

22. Henderson JM (2003) Human gaze control during real-world scene perception. Trends Cogn Sci 7(11):498–504

23. Henderson JM, Graham PL (2008) Eye movements during scene viewing: evidence for mixed control of fixation durations. Psychon Bull Rev 15(3):566–573

24. Itti L, Koch C (2000) A saliency-based search mechanism for overt and covert shifts of visual attention. Vis Res 40(10–12):1489–1506

25. Jagla F (2016) Saccadic eye movements as a marker of mental disorders. Physiol Res 65(Supplementum 3):S365-S371

26. Kapur S (2003) Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. Am J Psychiatry 160(1):13–23. https://doi.org/10.1176/appi.ajp.160.1.13

27. Kay SR, Fiszbein A, Opler LA (1987) The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 13(2):261–276

28. Lehrl S (2005) Mehrfachwahl-Wortschatz-Intelligenztest. MWT-B; Manual zum MWT-B, 5, unveränd. Aufl. Spitta-Verl., Balingen

29. Lencer R, Reilly JL, Harris MS et al (2010) Sensorimotor transformation deficits for smooth pursuit in first-episode affective psychoses and schizophrenia. Biol Psychiatry 67(3):217–223. https://doi.org/10.1016/j.biopsych.2009.08.005

30. Lencer R, Sprenger A, Reilly JL et al (2015) Pursuit eye movements as an intermediate phenotype across psychotic disorders: evidence from the B-SNIP study. Schizophr Res 169(1–3):326–333. https://doi.org/10.1016/j.schres.2015.09.032

31. Levy DL, Sereno AB, Gooding DC et al (2010) Eye tracking dysfunction in schizophrenia: characterization and pathophysiology. Curr Top Behav Neurosci 4:311–347

32. Loughland CM, Williams LM, Gordon E (2002) Visual scan-paths to positive and negative facial emotions in an outpatient schizophrenia sample. Schizophr Res 55(1–2):159–170. https://doi.org/10.1016/S0920-9964(01)00186-4

33. Lui S, Yao L, Xiao Y et al (2015) Resting-state brain function in schizophrenia and psychotic bipolar probands and their first-degree relatives. Psychol Med 45(1):97–108. https://doi.org/10.1017/S003329171400110X

34. Montgomery SA, Asberg M (1979) A new depression scale designed to be sensitive to change. Br J Psychiatry 134:382–389

35. Moritz S, Woodward TS (2007) Metacognitive training in schizophrenia: from basic research to knowledge translation and intervention. Curr Opin Psychiatry 20(6):619–625. https://doi.org/10.1097/YCO.0b013e3282f0b8ed

36. Nagel M, Sprenger A, Nitschke M et al (2007) Different extraretinal neuronal mechanisms of smooth pursuit eye movements in schizophrenia: an fMRI study. Neuroimage 34(1):300–309. https://doi.org/10.1016/j.neuroimage.2006.08.025

37. Nagel M, Sprenger A, Hohagen F et al (2008) Cortical mechanisms of retinal and extraretinal smooth pursuit eye movements to different target velocities. Neuroimage 41(2):483–492. https://doi.org/10.1016/j.neuroimage.2008.02.058

38. Nagel M, Sprenger A, Steinlechner S et al (2012) Altered velocity processing in schizophrenia during pursuit eye tracking. PLoS One 7(6):e38494. https://doi.org/10.1371/journal.pone.0038494

39. Phillips ML, David AS (1997) Visual scan paths are abnormal in deluded schizophrenics. Neuropsychologia 35(1):99–105. https://doi.org/10.1016/S0028-3932(96)00061-9

40. Phillips ML, David AS (1998) Abnormal visual scan paths. A psychophysiological marker of delusions in schizophrenia. Schizophr Res 29(3):235–245. https://doi.org/10.1016/S0920-9964(97)00097-2

41. Smith TJ, Mital PK (2013) Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. J Vis 13(8). https://doi.org/10.1167/13.8.16

42. Sprenger A, Friedrich M, Nagel M et al (2013) Advanced analysis of free visual exploration patterns in schizophrenia. Front Psychol 4:737. https://doi.org/10.3389/fpsyg.2013.00737

43. Startsev M, Lee A, Dorr M Optimizing clustering-based smooth pursuit detection. In: European conference on visual perception 2017

44. Steffens M, Becker B, Neumann C et al (2016) Effects of ketamine on brain function during smooth pursuit eye movements. Hum Brain Mapp 37(11):4047–4060. https://doi.org/10.1002/hbm.23294

45. Sweeney JA, Luna B, Haas GL et al (1999) Pursuit tracking impairments in schizophrenia and mood disorders: step-ramp studies with unmedicated patients. Biol Psychiatry 46(5):671–680

46. Tatler BW (2007) The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor biases and image feature distributions. J Vis 7(14): 4.1–17. https://doi.org/10.1167/7.14.4

47. Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. Nature 381(6582):520–522. https://doi.org/10.1038/381520a0

48. Trillenberg P, Sprenger A, Talamo S et al (2016) Visual and non-visual motion information processing during pursuit eye tracking in schizophrenia and bipolar disorder. Eur Arch Psychiatry Clin Neurosci. https://doi.org/10.1007/s00406-016-0671-z

49. Velichkovsky BM (2002) Heterarchy of cognition: the depths and the highs of a framework for memory research. Memory 10(5–6):405–419. https://doi.org/10.1080/09658210244000234

50. Verleger R, Talamo S, Simmer J et al (2013) Neurophysiological sensitivity to attentional overload in patients with psychotic disorders. Clin Neurophysiol 124(5):881–892. https://doi.org/10.1016/j.clinph.2012.11.003

51. Vig E, Dorr M, Cox D, Large-Scale Optimization of hierarchical features for saliency prediction in natural images. In: 2014 IEEE conference on computer vision and pattern recognition (CVPR), pp 2798–2805

52. Vig E, Dorr M, Martinetz T et al (2012) Intrinsic dimensionality predicts the saliency of natural dynamic scenes. IEEE Trans Pattern Anal Mach Intell 34(6):1080–1091. https://doi.org/10.1109/TPAMI.2011.198

53. Williams LM, Loughland CM, Gordon E et al (1999) Visual scanpaths in schizophrenia. Is there a deficit in face recognition? Schizophr Res 40(3):189–199. https://doi.org/10.1016/S0920-9964(99)00056-0

54. Wittchen HU, Zaudig M, Fydrich T (1997) Strukturiertes Klinisches Interview für DSM-IV, Achse I und II (SKID) [Structured Clinical Interview for DSM-IV Axis I and II Disorders (SCID)]. Hogrefe, Göttingen

55. Wolfe JM (1994) Guided search 2.0 a revised model of visual search. Psychon Bull Rev 1(2):202–238. https://doi.org/10.3758/BF03200774

56. Wylie GR, Clark EA, Butler PD et al (2010) Schizophrenia patients show task switching deficits consistent with N-methyl-D-aspartate system dysfunction but not global executive deficits: implications for pathophysiology of executive dysfunction in schizophrenia. Schizophr Bull 36(3):585–594. https://doi.org/10.1093/schbul/sbn119

57. Zelinsky GJ (2012) TAM: explaining off-object fixations and central fixation tendencies as effects of population averaging during search. Vis cogn 20(4–5):515–545. https://doi.org/10.1080/13506285.2012.666577