Diagnostic Accuracy of Eye Movements in Assessing Pedophilia

Peter Fromberger, MSc,* Kirsten Jordan, PhD,* Henrike Steinkrauss, stud. med.,* Jakob von Herder, stud. med.,* Joachim Witzel, MD,† Georg Stolpmann, MD,* Birgit Kröner-Herwig, PhD,‡ and Jürgen Leo Müller, MD*

*Department for Forensic Psychiatry and Psychotherapy, Georg-August-University of Göttingen, Göttingen, Germany; †Central State Forensic Psychiatric Hospital of Saxony-Anhalt, Stendal, Germany; ‡Clinical Psychology and Psychotherapy, Georg-August-University of Göttingen, Göttingen, Germany

DOI: 10.1111/j.1743-6109.2012.02754.x

ABSTRACT

Introduction. Given that recurrent sexual interest in prepubescent children is one of the strongest single predictors for pedosexual offense recidivism, valid and reliable diagnosis of pedophilia is of particular importance. Nevertheless, current assessment methods still fail to fulfill psychometric quality criteria.

Aims. The aim of the study was to evaluate the diagnostic accuracy of eye-movement parameters in regard to pedophilic sexual preferences.

Method. Eye movements were measured while 22 pedophiles (according to ICD-10 F65.4 diagnosis), 8 non-pedophilic forensic controls, and 52 healthy controls simultaneously viewed the picture of a child and the picture of an adult. Fixation latency was assessed as a parameter for automatic attentional processes and relative fixation time to account for controlled attentional processes.

Main Outcome Measures. Receiver operating characteristic (ROC) analyses, which are based on calculated age-preference indices, were carried out to determine the classifier performance. Cross-validation using the leave-one-out method was used to test the validity of classifiers.

Results. Pedophiles showed significantly shorter fixation latencies and significantly longer relative fixation times for child stimuli than either of the control groups. Classifier performance analysis revealed an area under the curve (AUC) = 0.902 for fixation latency and an AUC = 0.828 for relative fixation time. The eye-tracking method based on fixation latency discriminated between pedophiles and non-pedophiles with a sensitivity of 86.4% and a specificity of 90.0%. Cross-validation demonstrated good validity of eye-movement parameters.


Key Words. Pedophilia; Eye Movements; Diagnostic Accuracy; Attention; Automatic Processes; Assessment; Pedophilic Sexual Preferences

Introduction

Sexual child abuse is a general risk factor for psychopathology in the growing child [1]. It is estimated that only 40–50% of all child abusers fulfill the diagnostic criteria for pedophilia, i.e., primarily the existence of pedophilic sexual interests [2]. According to the American Psychiatric Association, pedophilia is defined as a recurrent sexual interest in prepubescent children characterized by persistent thoughts, fantasies, urges, sexual arousal, or behavior [3]. The strongest single predictor for sexual-offense recidivism is the existence of deviant sexual preferences [4]. This fact highlights the importance of valid and reliable diagnosis of deviant pedophilic preferences. Nevertheless, a lack of diagnostic accuracy is one major problem in treatment and risk assessment of pedophilia,
Diagnostic Accuracy of Eye Movements in Assessing Pedophilia

Diagnostic Accuracy of Eye Movements in Assessing Pedophilia

Assessment of Deviant Pedophilic Preferences

Self-reports are the diagnostic standard in Western Europe, but demonstrate insufficient validity and reliability, especially due to the tendency of participants to answer in a socially desirable manner [6]. Penis plethysmography (PPG) is the gold standard for assessing pedophilic interests in North America and has demonstrated good classification accuracy [7–9]. Nevertheless, the application of PPG is limited by certain factors. Phallometric testing is very intrusive [5]; it has been criticized for its high proportion of non-responders [10], and it is still controversial in terms of discriminant validity and selectivity [11,12].

Recent methods have been developed to assess sexual pedophilic interests by cognitive approaches [13]. One of the most elaborate approaches is the Implicit Association Test (IAT). Originally developed to identify implicit racist attitudes, the IAT was adapted to measure implicit child-sex associations to discriminate pedophiles from non-pedophiles. Furthermore, the IAT differentiates between pedophiles and hebephiles by demonstrating that only pedophiles exhibit implicit child-sex associations [14]. Although initial psychometric data revealed good reliability and consistency, the IAT seems prone to faking behavior [15]. Another critical point may be that the existence of implicit child-sex associations, as measured by the IAT, is neither a necessary nor a sufficient condition for the diagnosis of pedophilia. Thus, the relationship between implicit associations and sexual arousal remains unclear [16]. Another indirect approach known since 1942 is viewing time (VT). This method is based on the finding that pedophiles look at deviant sexual stimuli significantly longer than nondeviant stimuli. In other words, participants direct more attention to sexually relevant stimuli than to sexually nonrelevant stimuli. In most cases, the VT measurement is combined with a measure of self-reported sexual attraction. Several VT methods of measuring pedosexual interest are already on the market, e.g., the Abel Assessment for Sexual Interest-2 (AASI; Abel Screening Inc., Atlanta, GA, USA; [17]). The VT approach demonstrated good sensitivity and specificity—in some studies better than PPG. However, in clinical settings its application is restricted by questionable reliability and validity [18,19]. Attention also plays a crucial role in other approaches which have not passed the experimental stage: e.g., the Choice-Reaction-Time Task [20], the Rapid Serial Visual Presentation Test (RSVP) [21], or an adaptation of the Emotional Stroop Task for Sexual Offenders [22]. Taken together, indirect methods based on attentional processes may be promising tools to overcome the existing problems associated with the assessment of pedophilic sexual interests. Nevertheless, the exact mechanisms underlying the attentional processes that produced the differences between child molesters and non-child-molesters remain unknown. Furthermore, most studies did not differentiate between subgroups of child molesters, especially pedophiles and non-pedophilic child molesters, so it is worth questioning whether these approaches really measure pedophilic interests at all.

A new entry to assess pedophilic preferences accrues from eye-tracking research. The measurement of eye movements offers one possibility to directly explore attentional processes. Humans’ ability to identify fine details is limited to two degrees of central vision (the foveal region of the retina) [23]. This limitation of acuity of the human visual system makes it possible to identify the features most interesting to the viewing subject by eye movements: For scene perception, the relevant human eye movements can be divided into fixations and saccades. Saccades are rapid eye movements that switch the fovea from one stimulus to the next. Acquisition of information cannot take place in contrast, fixations are defined as a time periods in which the eye does not move (except for micro-saccades), and information is acquired [24]. The visual attention is directed to that part of the visual field on which the eye is fixating [25]. Furthermore, eye-tracking measures bias in initial orienting, as well as in the maintenance of attention [26,27].

Assessing Pedophilic Interests Using Eye Movements

Eye tracking has been applied since the end of the 19th century [28], but its relevance for sexuality has only been discovered quite recently. The vast majority of these studies focused on gender differences and demonstrated significant differences between male and female observers of erotic stimuli [29–33]. Other studies used eye tracking to explore attentional processes in the context of mating behavior from a more evolutionary perspective [34,35].

All these studies measured eye movements while adults looked at pictures or watched videos. Therefore, they do not provide relevant information for the assessment of pedophilic interests. Renaud et al. [36] first presented preliminary results concerning the potential of eye tracking to assess deviant sexual preferences. The authors used virtual-reality technology and immersed eight pedophile patients and eight nondeviant controls in a Cave Automatic Virtual Environment (CAVE)-type immersive system with 3D virtual characters. The virtual characters consisted of a 6-year-old child, an adult female, and a textureless virtual character. During the immersion, eye movements were assessed along with penile-plethysmographic data. The results demonstrated that pedophile subjects presented overall gaze-behavior dynamics of a lower fractal complexity (i.e., the gaze-behavior patterns of pedophiles consisted of fewer patterns that recur on finer and finer scales), as well as overall stronger relative erectile responses compared with normal subjects. However, numbers of fixations and fixation durations did not differ between the two groups. The authors could not decide whether the differences in fractal complexity were based on differences in low-level stimulus features or differences due to the sexual attractiveness of the persons. Although there were no significant differences between pedophiles and controls in classical eye-movement parameters, the study provides initial indications of the potential of using eye movements in assessing pedophilic interests.

Only two further eye-tracking studies used adult and child stimuli. Hall et al. [37] presented pictures of different-aged persons (babies, prepubescent children, and adults of different ages) to adult heterosexual males and females. Eye movements demonstrated that male participants directed significantly more attention to their sexually preferred pictures (20-year-old women) than to their non-preferred pictures. In contrast, female participants showed no age-specific effects. Based on this result, the authors concluded that eye tracking could be a sensitive tool for the assessment of sexual interests in human males, as well as in forensic contexts. Unfortunately, the eye-movement parameters in this study represented controlled attentional processes and are, therefore, susceptible to any attempt to falsify the results. One possibility to overcome this problem was demonstrated by Fromberger et al. [38]. We presented pairs of sexually preferred (images of women) and sexually non-preferred images (images of girls, boys, or men) to heterosexual males while eye movements were measured. Early attentional processing (initial orienting) was assessed by the number of first fixations, while late attentional processing (maintenance of attention) was assessed by relative fixation time. Results showed that relative fixation time was significantly longer for sexually preferred stimuli than for sexually non-preferred stimuli. Furthermore, the initial fixation was more often directed toward the preferred sexual stimulus. Our study showed for the first time an attentional bias in favor of sexually preferred stimuli when presented simultaneously with sexually non-preferred pictures. This initial orientation to sexually relevant stimuli has a high potential for the assessment of sexual interests because it seems to be an automatic attentional process and does not underlie conscious appraisal [38].

**Aims of the Study and Hypotheses**

We have already demonstrated that measuring eye movements can uncover both automatic and controlled attentional processes in healthy subjects while viewing sexually relevant and sexually non-relevant stimuli [38]. Cognitive approaches based on automatic processes while viewing sexual stimuli offer a promising alternative to the traditional assessment of pedophilic preferences, particularly as they are probably not susceptible to conscious attempts to falsify the results. Therefore, the aim of the study was to evaluate the diagnostic accuracy of eye-movement parameters in respect to pedophilic sexual preferences. Spiering and Everaerd’s theoretical model [39] proposes that sexually relevant features of a stimulus are pre-attentively selected and automatically trigger focal attention to these sexual aspects. Based on this and our recently published experimental design [38], we hypothesized: (i) that the relative fixation time for stimuli which are sexually relevant for pedophilic subjects (children) is significantly longer for pedophiles than for non-pedophiles; (ii) that the fixation latency for pictures of children is significantly shorter in the pedophilic group than in the non-pedophilic group; and (iii) that eye-movement variables would coincide with participants’ group status at a better-than-chance level.

**Methods**

**Participants**

Altogether, data of 22 pedophilic subjects, 8 forensic inpatients with no history of sexual assault
against children, and 52 healthy subjects were analyzed. Healthy subjects were recruited via a notice posted on bulletin boards in Göttingen and on inquiry at a police-officer school. Pedophilic subjects and forensic control subjects were recruited at high-security, forensic-psychiatric hospitals. Recruitment and study realization take place in 2010. Inclusion criteria for the pedophilic group were a cross-validated diagnosis of pedophilia (ICD-10 F65.4) by two experienced clinicians and mandatory hospitalization under treatment order for a child-sexual-abuse offense (validated through forensic records). Inclusion criteria for the forensic control group were the absence of a diagnosis of pedophilia (cross-validated by two experienced clinicians), no child-sexual-abuse offense, and mandatory hospitalization under treatment order for an adult-sexual-abuse offense (validated through forensic records). Inclusion criteria for the healthy, non-forensic control group were the absence of any psychiatric illnesses, deviant sexual fantasies or behavior (validated by an extensive psychiatric and sexual anamnesis conducted in a systematic oral interview about the case history of the subject by one experienced clinician). Exclusion criteria (especially for the inpatient groups) were an acute psychotic episode or substance abuse during the previous month, no agreement between the two clinicians in respect to the diagnosis of pedophilia, or incapability or refusal to sign informed consent. Due to these specifications and other conditions (e.g., no informed consent) 35 (3 pedophiles and 32 non-pedophiles) out of 65 screened forensic inpatients had to be excluded. None of the non-forensic healthy controls had to be excluded.

Table 1 summarizes the characteristics of the three subject groups with regard to sexual orientation, age, intelligence, hospitalization, and psychiatric diagnosis. Intelligence was assessed by the Wechsler Adult Intelligence Scale (German version) [40], and personality disorders were assessed by the Structured Clinical Interview for Diagnostic and Statistical Manual, fourth edition (DSM-IV) (German version) [41]. Assessment of sexual orientation for both the pedophilic group and the forensic control group was based on the victims’ gender. Sexual orientation of non-forensic controls was assessed by the Kinsey Scale [42], accepting only ratings from 0 to 1 (exclusively and predominantly heterosexual) or 5 to 6 (predominantly or exclusively homosexual). As shown in Table 1, the three groups were not homogenous with regard to their sexual orientation or their psychiatric diagnosis and differed significantly in respect to age, intelligence, and hospitalization. Pedophilic subjects demonstrated a median score of 5.00 (range 2.00–5.00) on the Screening Scale for Pedophilic Interests, which identified the group as a high-risk sample in respect to recidivism [43,44]. The pedophilic group had been convicted for sexually abusing an average of 6.05 children (standard error [SE] = 1.03, range 1–22). Child victims were on average 9.03 years of age (SE = 0.47 years, range 3.50–12.50 years). The minority of the pedophilic group were deniers (23%), indicating that the majority of the pedophilic subjects admitted their pedophilic fantasies or urges. Deniers and non-deniers only differed in their subjective self-evaluations in respect to pedophilic fantasies or urges; deniers did not admit their pedophilic fantasies or urges. The diagnosis for deniers and non-deniers was obtained from forensic records, their hospitalization history, and their sexual development and current sexual activities (sexual anamnesis).

All participants had normal visual acuity or corrected-to-normal visual acuity. All of them provided written informed consent before participating in the experiment. The study was approved by the Ethics Committee of the Medical Faculty of Georg-August University of Göttingen.

Procedure and Stimuli
Figure 1 shows the experimental design. Eye movements were recorded over the whole experiment (64 trials). Before each trial, a fixation cross (approximately 1° × 1°) appeared at the center of the screen. As soon as the participant had gazed at the fixation cross for at least 500 milliseconds, the next trial started automatically. This ensured that every participant would be looking at the middle of the stimulus display at the beginning of a trial. Then two pictures appeared and remained for 5,000 milliseconds. After each stimulus presentation, a question appeared (“Was one of these persons sexually more attractive?”), and the participant had to respond using the computer mouse. This task was introduced to distract the participant from eye-movement measures and had no content-related meaning. The participants were given the chance to rest after the first half of the experiment. At the beginning of each experiment and after the rest, a calibration was performed that consisted of instructing the participant to fixate on nine points on the display area.

Each stimulus display showed two pictures that were presented in two opposing corners of the computer screen (see Figure 1 for a stimulus example). Two different stimulus combinations
Table 1  Detailed characteristics of the subject groups. Shown are sexual orientations, demographic data, and ICD-10 diagnosis (n.a. = test statistic is not applicable).

<table>
<thead>
<tr>
<th></th>
<th>Pedophiles (N = 22)</th>
<th>Forensic controls (N = 8)</th>
<th>Non-forensic controls (N = 52)</th>
<th>Test statistic Overall group differences</th>
<th>P</th>
<th>Test statistic Pedophiles vs. forensic controls</th>
<th>P</th>
<th>Test statistic Pedophiles vs. non-forensic controls</th>
<th>P</th>
<th>Test statistic Forensic controls vs. non-forensic controls</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sexual orientation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterosexual</td>
<td>9 (40.9%)</td>
<td>9 (100%)</td>
<td>36 (69.2%)</td>
<td>(\chi^2 (df = 2)) 10.27</td>
<td>0.006</td>
<td>8.34</td>
<td>0.004</td>
<td>5.20</td>
<td>0.023</td>
<td>3.36</td>
<td>0.067</td>
</tr>
<tr>
<td>Homosexual</td>
<td>7 (31.8%)</td>
<td>0 (0%)</td>
<td>16 (30.8%)</td>
<td>(\chi^2 (df = 1)) 3.46</td>
<td>0.177</td>
<td>3.32</td>
<td>0.068</td>
<td>0.01</td>
<td>0.929</td>
<td>3.36</td>
<td>0.067</td>
</tr>
<tr>
<td>Bisexual</td>
<td>6 (27.3%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>(\chi^2 (df = 1)) 17.66</td>
<td>&lt;0.001</td>
<td>2.73</td>
<td>0.099</td>
<td>15.43</td>
<td>&lt;0.001</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td><strong>Demographic data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years (\pm 1\ SE)</td>
<td>42.06 (\pm 14.50)</td>
<td>36.68 (\pm 5.09)</td>
<td>25.27 (\pm 1.01)</td>
<td>(\chi^2 (df = 2)) 30.85</td>
<td>&lt;0.001</td>
<td>-0.89</td>
<td>0.185</td>
<td>5.41</td>
<td>&lt;0.001</td>
<td>2.35</td>
<td>&lt;0.009</td>
</tr>
<tr>
<td>Intelligence() (overall IQ (\pm 1\ SE)</td>
<td>76.52 (\pm 3.63)</td>
<td>76.88 (\pm 2.69)</td>
<td>117.58 (\pm 1.52)</td>
<td>(\chi^2 (df = 2)) 51.32</td>
<td>&lt;0.001</td>
<td>0.37</td>
<td>0.357</td>
<td>-6.32</td>
<td>&lt;0.001</td>
<td>-4.52</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Hospitalization() (month (\pm 1\ SE)</td>
<td>121.82 (\pm 14.50)</td>
<td>116.00 (\pm 42.57)</td>
<td>0 (0)</td>
<td>(\chi^2 (df = 2)) 72.97</td>
<td>&lt;0.001</td>
<td>-1.15</td>
<td>0.125</td>
<td>8.36</td>
<td>&lt;0.001</td>
<td>7.08</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>ICD-10 diagnosis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pedophilia (F65.4)</td>
<td>22 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>(\chi^2 (df = 2)) 82.00</td>
<td>&lt;0.001</td>
<td>30.00</td>
<td>&lt;0.001</td>
<td>74.00</td>
<td>&lt;0.001</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Organic mental disorder (F00–F09)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Substance abuse/dependence</td>
<td>9 (40.91%)</td>
<td>2 (25%)</td>
<td>0 (0%)</td>
<td>23.30</td>
<td>&lt;0.001</td>
<td>0.64</td>
<td>0.424</td>
<td>24.22</td>
<td>&lt;0.001</td>
<td>13.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Schizophrenia (F20–F29)**</td>
<td>3 (13.64%)</td>
<td>2 (25%)</td>
<td>0 (0%)</td>
<td>10.55</td>
<td>0.005</td>
<td>0.55</td>
<td>0.460</td>
<td>7.39</td>
<td>0.007</td>
<td>13.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Affective disorders (F30–F39)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Neurotic disorders (F40–F49)</td>
<td>2 (9.09%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>5.59</td>
<td>0.061</td>
<td>0.78</td>
<td>0.378</td>
<td>4.86</td>
<td>0.028</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Behavioral disorders with somatic disorders (F50–F59)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Personality disorders (F60–F69)**</td>
<td>9 (40.91%)</td>
<td>5 (62.50%)</td>
<td>0 (0%)</td>
<td>31.19</td>
<td>&lt;0.001</td>
<td>1.10</td>
<td>0.295</td>
<td>24.22</td>
<td>&lt;0.001</td>
<td>35.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mental disorders (F70–F79)</td>
<td>3 (13.64%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>8.49</td>
<td>0.014</td>
<td>1.21</td>
<td>0.271</td>
<td>7.39</td>
<td>0.007</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Developmental disorders (F80–F89)</td>
<td>1 (4.55%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>2.76</td>
<td>0.252</td>
<td>0.38</td>
<td>0.540</td>
<td>2.40</td>
<td>0.122</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Behavioral disorders with onset in childhood (F90–F99)</td>
<td>1 (4.55%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>2.76</td>
<td>0.252</td>
<td>0.38</td>
<td>0.540</td>
<td>2.40</td>
<td>0.122</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

*For the pedophilic group and the forensic control group, sexual orientation was assessed based on the victims' gender. Sexual orientation of non-forensic controls was assessed with the Kinsey scale [43], accepting only ratings from 0 to 1 (exclusively and predominantly heterosexual) or 5 to 6 (predominantly or exclusively homosexual).

\(\)Intelligence was assessed with the Wechsler Adult Intelligence Scale in the German version [41].

\(\)Hospitalization reflects the overall time duration of the subjects in forensic hospitals.

\(\)Participants with an F10–F19 ICD-10 diagnosis had no active substance abuse at least in the last month.

\(\)Participants with an F20–F29 ICD-10 diagnosis had no acute psychotic episode at least during the last month.

\(\)Personality disorders (without F65.4) were assessed with SKID-II [42]. Four pedophiles fulfilled the criteria for a borderline personality disorder, two for a narcissistic personality disorder, one for a histrionic personality disorder, three for an antisocial personality disorder, four for and avoidant personality disorder, and one for a paranoid personality disorder. One forensic control fulfilled the criteria for an avoidant personality disorder, two for an antisocial disorder, one for a schizotypal personality disorder, and one for a personality disorder not otherwise specified. Note that three pedophiles fulfilled the criteria for more than one personality disorder.

SE = standard error; df = degrees of freedom
exist, either the picture of a girl combined with the picture of a woman, or the picture of a boy combined with the picture of a man. The combination of pictures was pseudo-randomized so that each picture was presented twice, but in two different combinations and at two different locations on the screen. Furthermore, the locations of the pictures were balanced across trials, and their distances from each other were kept constant. Therefore, the spatial location of the picture is not expected to explain possible attentional effects between the stimuli. The height of all pictures was 412 pixels (which equals to 11.4° of visual angle at a viewing distance of 70 cm), with varying widths between 91 pixels (2.5°) and 280 pixels (7.8°). The picture pairs were matched in width. The distance between the two pictures was 16.4° (distance from center of the first picture to the center of the second picture). Thus, there was a minimal distance of 4° from the innermost border of one picture to the innermost border of the second picture.

Subsequent to the experimental phase, the participants rated all stimuli with respect to sexual arousal and valence on a nine-point Likert scale (1 = unpleasant/not arousing, 9 = pleasant/arousing). The time from stimulus onset to the completion of the second rating (sexual arousal rating) was measured without the participant’s knowledge to additionally assess viewing time.

All stimuli were selected from the Not-Real-People (NRP) Picture Set [45,46]. The NRP Picture Set contains a total of 160 colored pictures of nude and clothed male and female persons at five different stages of pubertal development based on Tanner’s categorization [47]. The Tanner’s categorization of the stimuli was assessed due to nine experts classifying the pictures into the Tanner stages (see [48] for more information). The pictures are not pornographic, depicting solely non-explicit sexual poses or sexual activities. In this study, only 64 nude males and females from Tanner stages 1, 2, 4, and 5 were used, four from each stage. To enlarge the usable stimulus set, each picture was mirrored so that eight female and eight male pictures of each Tanner stage were included. Pictures from Tanner stages 1 and 2 were combined into the category “child,” and pictures

Figure 1 Illustration of the time sequence of one trial. Note that these sample pictures were not among the experimental stimuli.
from Tanner stages 4 and 5 were combined into the category “adult.” The original NRP pictures had different-colored backgrounds, luminance levels, and complexity levels. Differences in visual low-level features, such as luminance and contrast, automatically attract attention in a bottom-up process, whereas the semantic content of a stimulus directs controlled attention in a top-down process [49]. To control confounding low-level features, all pictures were converted into grayscale pictures, and the colored backgrounds of all pictures were replaced by a consistent grayscale background. All pictures were preprocessed to adjust them in respect to luminance and contrast. The luminance level was assessed by converting the images into the Hue, Saturation, and Lightness (HSL) color space, reading out the luminance value, and calculating the mean luminance value for the entire image. The complexity of the pictures was assessed according to the number of bytes of the compressed image file size in JPEG format. Studies have demonstrated that the compressed image file size positively correlates with the image file complexity, as well as the human subjective judgment of picture complexity [50,51]. As luminance seems to influence the compression file size besides complexity [52], our preprocessing for luminance should control this factor. The stimulus categories did not differ significantly in respect to luminance, \( F(3, 60) = 1.04, \) and complexity, \( F(3, 60) = 0.51. \) Thus, low-level stimulus characteristics were probably not responsible for any possible attention effects between the stimulus categories.

**Apparatus**

Eye movements were measured using an SMI iView X RED eye tracker (SensoMotoric Instruments GmBH, Berlin, Germany) in combination with an iView X workstation by measuring the corneal reflection and dark pupil with a video-based infrared eye camera. The SMI RED system is a contact-free, remote-controlled eye tracking device with automatic eye and head tracker assuring that slight head movements (within a range of approximately 40 cm) are automatically compensated. It is not necessary to immobilize the head using a bite bar. However, we needed to ensure that the participants did not move out of the compensable range and asked them to rest their chins on their nondominant hands. The iView X RED system works with a spatial resolution of <0.1° of visual angle, a temporal resolution of 60 Hz and a gaze-position accuracy of <0.4° of visual angle. Stimuli were presented on a 19-inch thin film transistor (TFT) monitor (resolution 1,280 × 1,024 pixels) at a refresh rate of 75 Hz. The participants were seated in a quiet room facing the monitor at eye level at a viewing distance of 27.6 inches from the monitor.

**Eye-Movement Data Preparation and Measures**

To identify fixations and saccades, the raw eye-movement data were basically analyzed with BeGaze 3 (SensoMotoric Instruments GmBH, Berlin, Germany). This program works with a dispersion-threshold identification algorithm to identify fixations. If eye movements remained stable within a circular area of 1° of visual angle for at least 100 milliseconds, this was classified as a fixation [53]. To analyze visual attention to the different aspects of the stimulus display, we divided each stimulus display into two areas of interests (AOIs). Each entire picture (woman, man, boy, girl) equated to one AOI. Two eye-movement variables were measured. Relative fixation time was defined as the sum fixation duration of all fixations located in the space of the relevant AOI divided by the whole presentation time. Relative fixation time is a measure for the overall attention a specific AOI attracts [54]. In contrast, fixation latency reflects early attentional processes. Fixation latency was defined as the duration from stimulus onset to the first fixation within an AOI, representing a measure for initial orienting [55,56].

**Statistical Analyses**

All statistics were performed with SAS 9.2 (SAS Institute Inc., Cary, NC, USA). To control for the unequal distribution of general sexual orientations within the three groups, only the sexually relevant pictures were analyzed [20]—for heterosexual subjects, only pictures of females (girls and women) and for homosexual subjects, only pictures of males (boys and men). For bisexual subjects, pictures of both males and females were included in the analysis. This strategy resulted in two stimulus age categories: child and adult.

Differences between the groups in respect to sexual arousal and valence ratings were analyzed using Kruskal–Wallis Tests within each stimulus age category, and post hoc tests were performed by Wilcoxon Rank Sum Tests. Due to technical problems, rating data for two pedophiles, one forensic control, and four non-forensic controls could not be analyzed.

Eye-movement and viewing-time data were analyzed using mixed linear models (SAS 9.2 proc
mixed procedure using the Restricted Maximum Likelihood (REML) method). Calculation of the mixed model was performed according to the Littell et al. [57] approach. The best fitting covariance structure for each dependent variable was identified by the Akaike Information Criteria, indicating that the variance components covariance structure fitted best for eye-movement data, while the compound symmetry covariance structure revealed best fit for viewing-time data. Significant differences due to age, intelligence, and hospitalization among the three groups (Table 1) were expected to affect the statistical results. A Welch Test (due to unequal variances) did indeed reveal significant differences among the sizes of the AOIs (girl, woman, boy, man), \( F(3, 5,792.03) = 81.92, \quad P < 0.001 \). Therefore, we modeled age, IQ, hospitalization, and the AOI-size as additional factors within the linear mixed models to eliminate influences of the confounding variables.

Linear mixed models with subject group (pedophilic, forensic controls, non-forensic controls) as a fixed between-subject factor, stimulus age (child, adult) as a fixed within-subject factor, stimulus as a random effect within each subject, and subject as a repeated effect were conducted for both eye-movement variables. A linear mixed model with group as a fixed between-subject factor, stimulus age as a fixed within-subject factor, and subject as a repeated effect was conducted for viewing time. Viewing-time data for three pedophiles, one forensic control, and four non-forensic controls could not be analyzed because of technical difficulties. Post hoc pairwise comparisons were performed to further analyze significant interactions.

**Main Outcome Measures**

We performed receiver operating characteristic (ROC) analyses to test how well eye-motion parameters differentiate between pedophilic and non-pedophilic subjects. Classifier performance was measured by the area under the curve (AUC). An area of 1 represents a perfect test, and an area of 0.50 represents a test at chance level [58,59]. Since our main purpose was not primarily to differentiate among pedophiles, forensic controls, and non-forensic controls, we pooled the two control groups into a (new) non-pedophilic group. An Age Preference Index (API) was also calculated for each eye-movement variable. The API was defined as the difference between adult and child pictures [20]. The higher the API was for fixation latency, the shorter the fixation latency was for child stimuli. The lower the API was for relative fixation time, the higher the relative fixation time was for child stimuli. The lower the API was for viewing time, the higher the proportion of viewing time allotted to child pictures was. Since viewing time represents a widely accepted measurement for pedophilic sexual interests [19], we also calculated the API for viewing time and compared the AUC of viewing time with the AUC of eye-movement variables. Since the ROC analyses performance measures are obtained by self-prediction and are more optimistic than in real life, classifier performance was cross-validated using the leave-one-out method. This method was applied to estimate the classifier performance for previously unknown data sets by setting aside one observation of our data set, building the model on the rest of the data set and using this model to predict the left-out record. This process was repeated for each observation and provided information about the generalization ability of the classifier [60].

**Results**

**Rating Data**

Table 2 presents the medians of the sexual-arousal and valence rating as a function of subject group. The sexual-arousal ratings of the three groups differed significantly within the adult, \( \chi^2(2) = 12.49, \quad P = 0.001 \), and within the child stimulus category, \( \chi^2(2) = 16.79, \quad P < 0.001 \). Pedophiles demonstrated significantly higher sexual-arousal ratings for child stimuli than non-forensic controls, \( Z = 3.63, \quad P < 0.001, \quad r = 0.44 \), and forensic controls,

<table>
<thead>
<tr>
<th>Rating</th>
<th>Stimulus age</th>
<th>Pedophiles</th>
<th>Forensic controls</th>
<th>Non-forensic controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sexual arousal</td>
<td>Child</td>
<td>2.00 (1.00–8.00)</td>
<td>1.00 (1.00–1.00)</td>
<td>1.00 (1.00–7.00)</td>
</tr>
<tr>
<td></td>
<td>Adult</td>
<td>4.00 (1.00–9.00)</td>
<td>1.50 (1.00–5.00)</td>
<td>6.00 (1.00–9.00)</td>
</tr>
<tr>
<td>Valence</td>
<td>Child</td>
<td>2.00 (1.00–7.50)</td>
<td>1.00 (1.00–3.00)</td>
<td>2.50 (1.00–7.50)</td>
</tr>
<tr>
<td></td>
<td>Adult</td>
<td>6.00 (1.00–9.00)</td>
<td>4.00 (1.00–6.00)</td>
<td>6.50 (3.00–9.00)</td>
</tr>
</tbody>
</table>
Non-forensic controls revealed significantly higher sexual arousal ratings for adult stimuli than forensic controls, Z = -3.63, P < 0.001, r = 0.49. All other post hoc tests revealed no significant difference. Valence ratings of the three groups also differed significantly within the adult, χ²(2) = 9.98, P = 0.003, and within the child stimulus category, χ²(2) = 7.15, P = 0.023. Pedophiles rated child and adult stimuli on the valence scale significantly higher than forensic controls (child: Z = -2.09, P = 0.016, r = 0.40; adult: Z = -1.96, P = 0.021, r = 0.38). Non-forensic controls rated both stimulus categories on the valence scale significantly higher than forensic controls (child: Z = -2.64, P = 0.002, r = 0.32; adult: Z = -3.26, P < 0.001, r = 0.40). In contrast, pedophiles did not differ significantly from non-forensic controls in terms of valence ratings of child or adult stimuli.

**Viewing Time**

Figure 2C shows the means and SEs for viewing time as a function of stimulus age category and group. The linear mixed model revealed a significant main effect for stimulus age, F(1, 71) = 40.68, P < 0.001, r = 0.64, and a significant group x stimulus age interaction, F(2, 71) = 14.60, P < 0.001. Only the age of the subjects had a significant influence on viewing time, F(1, 68) = 9.59, P = 0.002, r = 0.12. The older the subject, the longer his viewing time. Pairwise comparisons demonstrated that pedophiles showed a significantly longer viewing time for child stimuli in comparison with forensic controls, t(71) = 2.13, P = 0.037, r = 0.25. No other pairwise comparison revealed significant differences.

**Fixation Latency**

Figure 2A shows the means and SEs for fixation latency as a function of stimulus age category and group. The linear mixed model revealed a significant main effect for the stimulus age category, F(1, 77) = 14.72, P < 0.001, r = 0.16, and a significant group x stimulus age interaction, F(2, 77) = 34.44, P < 0.001. The age of the subject, intelligence, hospitalization, and AOI-size had no significant influence. Pairwise comparisons demonstrated significantly shorter fixation latencies for child stimuli among the pedophiles than among forensic controls, t(77) = 3.19, P = 0.002, r = 0.34, and non-forensic controls, t(77) = 2.29, P = 0.025, r = 0.25. Therefore, our hypothesis that is the fixation latency for pictures showing children is significantly shorter in the pedophilic group than in the non-pedophilic group was confirmed. Fixation latencies for adult stimuli did not differ significantly between the groups.
Relative Fixation Time

Figure 2B shows the means and SEs for relative fixation time as a function of stimulus age category and group. The linear mixed model revealed a significant main effect for the stimulus age category, $F(1, 77) = 1,253.62$, $P < 0.001$, $r = 0.94$, and a significant group x stimulus age interaction, $F(2, 77) = 170.98$, $P < 0.001$. The size of the AOIs had a small, but significant, influence on the relative fixation time, $F(1, 5,354) = 36.02$, $P < 0.001$, $r = 0.01$. The larger the AOI was, the shorter the relative fixation time was. Subject age, intelligence, and hospitalization had no significant influence. Pairwise comparisons showed that pedophiles had a significantly longer relative fixation time for child stimuli than non-forensic controls, $t(77) = -2.97$, $P = 0.039$, $r = 0.32$, and forensic controls, $t(77) = 6.01$, $P < 0.001$, $r = 0.57$. The hypothesis that the relative fixation time for stimuli which are sexually relevant to pedophilic subjects (children) is significantly longer for pedophiles than for non-pedophiles was confirmed. In contrast, pedophiles demonstrated a significantly shorter relative fixation time for adult stimuli than non-forensic controls, $t(77) = 3.15$, $P = 0.002$, $r = 0.34$, and forensic controls, $t(77) = 2.30$, $P = 0.024$, $r = 0.25$. Non-forensic controls and forensic controls did not differ significantly in their relative fixation time for adult or for child stimuli.

ROC Analyses

Table 3 summarizes the results of the receiver operating characteristic analyses. Figure 3 shows the ROC curves and dot diagrams for fixation latency, relative fixation time, and viewing time. The API for fixation-latency data, the API for relative fixation time, and the API for viewing time demonstrated a significant difference between pedophiles and non-pedophiles. Thus, all three variables could be useful to identify the subject status.

ROC analyses revealed that fixation-latency measurement discriminated between pedophiles and non-pedophiles (AUC = 0.902) with highest accuracy. Using a cutoff criterion of $-33.676$ milliseconds (an API score above $-33.676$ milliseconds as an indication for pedophilia), the API based on fixation latency differed between pedophiles and non-pedophiles with a sensitivity (probability that a pedophile will be correctly classified as pedophilic) of 86.4% and a specificity (probability that a non-pedophile will be correctly classified as non-pedophilic) of 90.0%. ROC analysis for relative fixation time revealed a somewhat lower AUC = 0.828 and for viewing time, an AUC = 0.759. From a statistical point of view, the AUC based on fixation latency did not differentiate significantly better between pedophiles and non-pedophiles than the AUC based on relative fixation time, $Z = 0.43$, $P = 0.668$, $r = 0.05$, or the AUC based on viewing time, $Z = 1.87$, $P = 0.062$, $r = 0.22$.

Cross-validation of the classifier performance using the leave-one-out method revealed an AUC = 0.891 for fixation latency, an AUC = 0.812 for relative fixation time, and an AUC = 0.726 for viewing time (Table 3).

Table 3 Overview over the Age Preference Index (API) and the Receiver Operating Characteristic (ROC)

<table>
<thead>
<tr>
<th></th>
<th>Fixation latency</th>
<th>Relative fixation time</th>
<th>Viewing time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>API</strong> Pedophiles (mean ± 1 SE)</td>
<td>111.17 ms ± 45.65 ms</td>
<td>8.88% ± 3.95%</td>
<td>225.25 ms ± 381.71 ms</td>
</tr>
<tr>
<td>Non-pedophiles (mean ± 1 SE)</td>
<td>-259.74 ms ± 32.57 ms</td>
<td>30.34% ± 1.59%</td>
<td>1,838.29 ms ± 265.75 ms</td>
</tr>
<tr>
<td>Test statistic</td>
<td>$t(80) = -6.13$</td>
<td>$t(28.12) = 5.01^*$</td>
<td>$t(73) = 3.24$</td>
</tr>
<tr>
<td>$P$</td>
<td>&lt;0.001</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>$R$</td>
<td>0.57</td>
<td>0.56</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>ROC</strong> AUC</td>
<td>0.902</td>
<td>0.828</td>
<td>0.759</td>
</tr>
<tr>
<td>$P$</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>95%-CI</td>
<td>0.808–0.977</td>
<td>0.709–0.947</td>
<td>0.633–0.885</td>
</tr>
<tr>
<td>Cutoff</td>
<td>≥33.68 ms</td>
<td>≥20.41%</td>
<td>≥904.88 ms</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>86.40%</td>
<td>81.82%</td>
<td>80.00%</td>
</tr>
<tr>
<td>Specificity</td>
<td>90.00%</td>
<td>80.00%</td>
<td>76.36%</td>
</tr>
<tr>
<td><strong>Cross-validation</strong> AUC</td>
<td>0.891</td>
<td>0.812</td>
<td>0.726</td>
</tr>
<tr>
<td>$P$</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>95%-CI</td>
<td>0.800–0.971</td>
<td>0.686–0.938</td>
<td>0.592–0.861</td>
</tr>
</tbody>
</table>

*Satterthwaite approximation due to unequal variances.

The API was defined as the difference between adult and child stimuli. The Area under the Curve (AUC) is measure for test accuracy. An area of 1 represents a perfect test; an area of .5 represents a test at chance level. Sensitivity is defined as the probability that a pedophilic subject will be correctly classified as pedophilic. Specificity is defined as the probability that a non-pedophilic subject will be correctly classified as non-pedophilic. Cross-validation was performed with the leave-one-out method.

SE = standard error; CI = confidence interval; ms = milliseconds

The aim of the study was to evaluate the diagnostic accuracy of eye-movement parameters in respect to pedophilic sexual preferences. The reported results demonstrated that the eye-tracking method enables a differentiation between pedophiles and non-pedophiles with high sensitivity and high specificity. An age-preference index based on fixation latency correctly classified 86.4% of the pedophiles with 10% false positives. Cross-validation demonstrated high validity of the classifier, since the AUC decreased only from an AUC = 0.902 to a cross-validated AUC = 0.891. Interestingly, on a descriptive level, the accuracy of the API based on eye movements that reflect automatic attentional processes (fixation latency) differentiated better than the API based on eye movements that reflect controlled attentional processes (relative fixation time). That could be an important result, especially in forensic settings: Automatic attentional processes are probably more difficult to fake than controlled attentional processes, which are consciously accessible. Therefore, the API based on fixation latency seems more suitable for assessing pedophilic sexual interests than the API based on relative fixation time. Moreover, relative fixation time is a measurement with a long time period (5 seconds). In contrast, fixation latency is a measurement with a very short time period. This difference between the two eye-movement variables already constitutes the advantage of fixation latency over relative fixation time. The reason for the lack of a statistically significant difference in regard to the accuracy between the two APIs could lie in the fact that the majority of the subjects were
Diagnostic Accuracy of Eye Movements in Assessing Pedophilia

1879

assumed severity in faking the fixation latency as an indicator of early attentional processes. The diminished, as fixation latency is considered to be faking or socially desirable behavior might be simple. Furthermore, it is less intrusive than the PPG. Another potential advantage over PPG is that the eye-tracking method reveals comparable or even better accuracy. Viewing time could correctly classify pedophiles in a range between 65.6% and 96.0% with false positives ranging between 23.0% and 35.3% [17]. As in comparison with the IAT, the eye-tracking method seems to provide better accuracy. PPG demonstrated sensitivity values between 68.0% and 88.6% and specificity values between 90% and 96.9% [7,9]. These studies examined child molesters in the broader sense, so that not exclusively pedophiles as defined in the DSM-IV were included. When examining only pedophiles, PPG demonstrated a somewhat lower sensitivity of 55% and specificity of 96.5% [8].

In summary, our results suggest that the eye-tracking method provides comparable or even better accuracy in the classification of pedophiles than existing methods that are, to some extent, already established in clinical settings. However, diagnostic accuracy is only one attribute a clinical test must fulfill. Another important aspect is the clinical usefulness [61]. In this context, we believe that the eye-tracking method offers certain advantages compared with existing methods. Due to a relatively short duration of the entire eye-tracking procedure (about 15 minutes) and the easy transportability and flexibility of the equipment (comparable to a standard personal computer), implementation in everyday clinical practice would be simple. Furthermore, it is less intrusive than the PPG. Another potential advantage over PPG is that faking or socially desirable behavior might be diminished, as fixation latency is considered to be an indicator of early attentional processes. The assumed severity in faking the fixation latency derives from eye-tracking studies while viewing emotional pictures. For example, Nummenmaa et al. [27] demonstrated that participants generally looked at highly emotional pictures (which were presented simultaneously with emotionally neutral pictures) first, even when the subjects were instructed to avoid looking at the emotional pictures. The authors concluded that these early attentional processes measured by eye tracking cannot be consciously influenced. Furthermore, from a more theoretical point of view, Spiering and Everaerd [39] assumed that sexual stimuli lead to an initial orienting to sexually relevant aspects. This is a well-known effect of other evolutionarily relevant stimuli, like spiders or snakes. The initial orienting is assumed to not be consciously influenceable because it is an automatic process. We already demonstrated that it is possible to measure the initial orienting to sexual stimuli with the eye-tracking method [38]. From the assumed automaticity of the initial orienting to sexually relevant stimuli, one can deduce that the eye-movement parameters reflecting it (like the fixation latency) can hardly be consciously influenced. This should, however, be tested directly in the context of pedosexual assessment in future research.

Some limitations have to be made. First, pedophiles and non-forensic controls were not matched in regard to age, intelligence, and hospitalization. It is unclear if pedophilia is directly linked to lower intellectual functioning, since the literature provided depended on the studied sample different results (e.g., [62] [63] [64]). We tried to overcome this limitation by including these potentially confounding variables as additional factors in the model. However, no significant influence on the eye-movement data was found. The forensic controls and the pedophiles were matched in regard to these confounding variables. The fact that forensic controls demonstrated results comparable to those of the non-forensic controls on both eye-movement parameters supports the absence of a significant influence of age, intelligence, and hospitalization. Second, the sample size, especially for forensic controls, was relatively low. Thus, any generalization of the results should only be made with caution. Third, sexual orientation was assessed with the Kinsey Scale in the non-forensic control group; in the pedophilic and in the forensic control group victims’ gender was the determining factor. This approach is not perfect, but it appears necessary due to socially desirable faking tendencies of forensic inpatients. Furthermore, a correlation analysis demonstrated a significant and
highly positive correlation between our group assignment based on victim’s gender and the self-evaluation of sexual orientation on the Kinsey Scale \( r = 0.70, P < 0.001 \). Fourth, the majority of pedophiles in our sample were non-deniers. Demonstrating that non-deniers can correctly be classified with our method is an initial, but important, step in establishing a new assessment method. One can deduce that a new assessment method really is a valid method only with non-deniers, but a new method must also be able to correctly classify deniers for clinical usage. Fifth, a faking test is necessary to demonstrate to what degree the eye-movement variables can be consciously influenced and are, therefore, susceptible to faking. Sixth, the possibility of cross contamination between the two groups existed because it was not possible to assure that some subjects in the non-pedophilic group did not suffer at least to some degree from pedophilic interests or fantasies with 100% reliability. We tried to exclude this with extended anamneses by experienced clinicians, but cross contamination was theoretically possible. Therefore, the additional application of PPG would be important in future research, as it is still the gold standard. In comparison with other assessment methods, our eye-tracking method is more expensive and requires specialized training for the clinicians. However, technical development in the area of eye-tracking is progressing rapidly, and prices are decreasing. Furthermore, we are working on more user-friendly software to enhance the usability of the method without specialized knowledge.

Conclusions

Our results demonstrated that measuring eye movements seems to be a promising approach to assess deviant pedophilic interests. In comparison with existing measurements, it provides comparable or better accuracy and, therefore, could improve the diagnostic validity of the assessment of pedophilia. Cross-validation indicated the acceptable validity of eye-movement variables.

Future research may concentrate on a combination of eye tracking with existing methods, such as viewing time or PPG, to obtain further insights into the mechanisms of these approaches or to better identify attempts at faking [65]. Valid and objective assessment methods, like our eye-tracking method, may also be helpful for therapy evaluation because their usage will heighten the reliability of the initial diagnosis [66]. Furthermore, it could be possible to measure the therapeutic outcome with the eye-tracking method. Therefore, future research should also concentrate on the ability of the eye-tracking method to measure therapy outcome and progress. Apart from our promising results and considering that a diagnosis of pedophilia is usually associated with enormous social consequences for the people involved, exhaustive future research must be performed, especially in respect to the psychometric criteria, before this new method should be used in clinical settings. Furthermore, it is important to emphasize that the diagnosis of deviant pedosexual preferences should not only be based on assessment methods like the eye-tracking, but should also always be combined with other information (e.g., offense history if available, detailed sexual anamnesis) to meet the complexity and requirements of diagnosing deviant sexual preferences.

Corresponding Author: Peter Fromberger, MSc, Department for Forensic Psychiatry and Psychotherapy, Georg-August-University of Göttingen, Rosdorfer-Weg 70, 37081 Göttingen, Germany. Tel: +49 (0)551 4022108; Fax: +49 (0)551 4022110; E-mail: peter.fromberger@medizin.uni-goettingen.de

Conflict of Interest: None.

Statement of Authorship

Category 1

(a) Conception and Design
Peter Fromberger; Kirsten Jordan; Jürgen Leo Müller

(b) Acquisition of Data
Peter Fromberger; Kirsten Jordan; Henrike Steinkrauss; Jakob von Herder; Joachim Witzel

(c) Analysis and Interpretation of Data
Peter Fromberger; Kirsten Jordan; Birgit Kröner-Herwig; Jürgen Leo Müller

Category 2

(a) Drafting the Article
Peter Fromberger

(b) Revising It for Intellectual Content
Peter Fromberger; Kirsten Jordan; Jürgen Leo Müller; Georg Stolpmann; Henrike Steinkrauss; Jakob von Herder; Joachim Witzel; Birgit Kröner-Herwig

Category 3

(a) Final Approval of the Completed Article
Peter Fromberger; Kirsten Jordan; Georg Stolpmann; Jakob von Herder; Henrike Steinkrauss; Birgit Kröner-Herwig; Jürgen Leo Müller; Joachim Witzel
Diagnostic Accuracy of Eye Movements in Assessing Pedophilia

References


