NEUROLOGY, PSYCHIATRY AND BRAIN RESEARCH 18 (2012) 1–7

Available at www.sciencedirect.com

SciVerse ScienceDirect

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

Improving seizure recognition by visual reinforcement

Andras Fogarasi [a,*], Gabor Szirtes [b], Gergo Zsiák [b], Andras Lorincz [b]

[a] Epilepsy Centre, Bethesda Children's Hospital, Budapest, Hungary
[b] Department of Software Technology and Methodology, Eötvös Loránd University, Budapest, Hungary

ARTICLE INFO

Article history:
Received 23 May 2011
Received in revised form 7 October 2011
Accepted 24 October 2011
Available online 22 November 2011

Keywords:
Epilepsy
Diagnosis
Seizure semiology
Eye-tracking
Education

ABSTRACT

Objective: Exact seizure description is an important diagnostic tool of epilepsy; however, epileptologists frequently get misleading data from lay eyewitnesses. The aim of this study is to find new methods for efficient training of seizure observation.

Methods: Twelve video-recorded seizures were observed by four groups of subjects with different expertise level in epilepsy: (1) *naïve observers*, (2) *pediatric residents*, (3) *epilepsy nurses*, and (4) *experts in video-EEG monitoring*. In half of the experiments, relevant parts of seizures were highlighted to direct the subjects' gaze during observations. Eye motion data of the observers were recorded by a Tobii T120 Eye Tracking system. After watching a seizure, subjects were asked to fill out a questionnaire regarding semiology data. To reinforce associations between verbal and visual information, subjects watched again the same clip. Observation skill improvement was accessed by comparing questionnaire scores and eye fixation data.

Results: Questionnaire mean scores increased and deviations decreased through the study demonstrating overall improvement for each group. Fixation time values were shorter in more experienced observers, however, this difference diminished by time reflecting the effect of the learning process. The chosen method for visual highlighting did not facilitate learning.

Conclusions: Our pilot study – using eye-tracking assessment in clinical epileptology for the first time – shows that seizure observational skills can be significantly facilitated by a combined use of video examples and questionnaires. With the increasing popularity of eye-tracking methods in applied sciences, our results may lead to novel approaches in epilepsy education of both lay subjects and professionals.

1. Introduction

Exact seizure description is an important tool of the diagnostic process on seizure- and epilepsy syndrome classification as well as during presurgical evaluation.[1] However, epileptologists frequently get misleading or partial data from eyewitnesses of seizures.[2] Our goal was to develop fast, but reliable methods by which subjects (e.g., care givers) can be efficiently trained to better recognize and describe epileptic seizures. Speed is an important factor as naïve subjects such as parents, teachers or care givers usually have limited time to take part in such training. Ideally a short learning phase should be immediately followed by a fast testing and evaluation phase. Earlier studies on learning demonstrate visual modality has a strong impact on the speed and success of learning in terms of retrieval and use of the gained knowl-

* Corresponding author. Address: Epilepsy Centre, Bethesda Children's Hospital, H-1146-Budapest, Bethesda Street 3, Hungary. Tel.: +36 70 7756884; fax: +36 1 3649070.
E-mail address: fogarasi@bethesda.hu (A. Fogarasi).

0941-9500/$ - see front matter
doi:10.1016/j.npbr.2011.10.002

edge.[3,4] In turn, we designed a method in which conventional tests and visual examples of seizures of different type were used in combination: the explanatory texts and questions (in the form of questionnaires about seizure semiology) provided a contextual frame of reference in which the visual example can be better verbalized.[5]

The main goal of the experiments was to see if subjects with varying knowledge about seizure semiology can be trained to provide better report on seizure events by the combined use of verbal questionnaires and video samples. In addition to the assessment of the questionnaires, eye movements recorded during watching were also analyzed. The rationale was that gaze as a primary behavioral response during information processing is shown in many cases to correlate with expertise level.[6,7] In what follows we describe the method and the pilot studies for accessing the viability of the method. We conclude by summarizing the results and pointing to future improvements and applications.

2. Material and methods

2.1. Seizure video samples

Twelve digital video records of different seizures were chosen by the reference from a pool of ca. 3000 seizures of 464 children who underwent long-term video-EEG monitoring at the Bethesda Children's Hospital (Budapest) between 2001 and 2009. Length of clips was 23–157 (mean: 76) seconds. Video samples were ranked by their complexity as judged by the reference. To facilitate the associations between the verbalized information items and the visual examples, in some of the experiments relevant parts of the visual space were highlighted on each frame according to the fixation data of the reference (highlighting was achieved by increasing the illumination on areas with higher fixation density; see Fig. 1). Observers were requested to review the video samples in three sessions (one session per week, in three consecutive weeks), each session consisting of four records. Seizure samples were grouped by their complexity (based on the evaluation of the reference) so session #1 included the least complex video samples, while session #3 contained the most complex ones. Seizures with the more compartments (e.g., seizure semiology components, lateralizing signs, vegetative signs) were the more sever. The study was approved by the local Ethics Committee of the Bethesda Children's Hospital.

2.2. Seizure observer groups

Observers were grouped on the basis of their expertise level on epilepsy. Group #1 contained eight naïve observers (informatics and technical university students) without any experience on epilepsy. Group #2 contained eight pediatric residents who already learned about epilepsy but had no practical experience viewing seizures. Group #3 contained eight nurses working in our Epilepsy Centre. They frequently see epileptic seizures on the ward but do not regularly watch or analyze video-recorded samples. Finally, group #4 contained six experts in seizure semiology (four epileptologists and two EEG assistants who regularly analyze epileptic seizures in the Video-EEG Unit. Reference data (both questionnaire responses and fixation data) were provided by pediatric epileptologist A.F. who has analyzed about 5000 video-recorded seizures in the past 10 years.

2.3. Eye-motion recording

During seizure watching, projected eye motions (in forms of xy coordinates on the display showing the stimuli) as well as fixation times of all observers were recorded by a Tobii T120 Eye Tracking system (Tobii Technology, Sweden; accuracy: 0.5°; drift: <0.3°; sampling rate: 120 Hz; type of display: 17" TFT, 1280×1024 pixels; distance from the display: $\sim$70 cm). The system measured the changes of the pupils' infrared reflection as a function of eye motion.

2.4. Questionnaire

After watching the video sample, observers were asked to fill out a questionnaire containing seizure semiology data regarding different aspects of the observed seizure (the whole Questionnaire can be found as Supporting Information). Besides semiological seizure classification data,[8] the questionnaires

Fig. 1 – (A) A sample frame of a video sample on a versive seizure. (B) Highlighted version of the same frame based on the attention map of the reference. Note that reference focused on the ictal features of patient's hands (right hand dystonia and left hand automatism can help lateralizing the seizure onset zone to the left hemisphere in this case).

included motor, consciousness, autonomic, and emotional axes of each seizure[9] as well as lateralizing signs[10] and post-ictal behavior.[11] Because of naïve observers' participation in the study, a short explanation was given after each semiological terminology on the questionnaire (e.g., "clonic seizure" = rhythmic jerking of a certain part of the body, face or extremities).

Having filled out the questionnaire, observers read the reference's answers. To facilitate the learning process, subjects were asked to watch the same video sample again. We also wanted to test if the learning effect can be amplified by highlighting the most relevant regions of patients during seizure (defined by large spatial densities of the fixation points extracted from the reference's scan path). For this reason, half of the observers in each group watched the highlighted version of the same sample at the post-test phase.

2.5. *Questionnaire scoring*

To analyze the written responses, we compared subjects' performances by extracting the Balanced Classification Rate[12] from the confusion matrix: BCR = 0.5 × (Sensitivity + Specificity), where Sensitivity (or Recall) = TP/(TP + FN) and Specificity = TN/(TN + FP) (TP, FP, TN and FN are true positive rate, false positive rate, true negative rate and false negative rate, respectively). The ground truth, for which the responses were compared against, was given by the reference.

2.6. *Analysis of eye tracking data*

Eye movements were analyzed in terms of fixations (pauses over informative regions of interest) and saccades (rapid movements between fixations). In defining fixation we followed the so called *velocity threshold identification method*.[13] First, the instantaneous velocity is calculated at every sample point as the distance in the xy coordinates measured in the current and previous sampling. Each sample points is then classified as fixation or saccade based on a simple velocity threshold (6 pixels/ms). Finally we merged the consecutive fixation points into a fixation group and calculated the fixation duration as the number of fixation points within the group times the sampling interval. Similarly to earlier eye fixation studies[13,14] we also established a minimal (100 ms) and maximal (2000 ms) fixation duration and any data outside this range were excluded from further analysis. An important difference between our setting and most other studies using eye tracking is that we applied moving stimuli as opposed to static images. Although there are some reports on using video clips[6,15], there is no validated method to follow. Although the device used in our study was expected to provide reliable data, we had to exclude recordings of observers wearing eyeglasses (as glasses disturbed the infra-red reflection properties of the pupils). Final group sizes for eye-tracking analysis were: six naïve subjects, seven residents, seven nurses and four experts.

2.7. *Statistics*

Improvements of BCR scores averaged over the subjects and video samples were measured by two sample (unpaired) one-sided *t*-test. Fixation duration distributions of all observers were visualized by histograms with 50 ms bins. To quantify differences between these distributions, non-parametric two-sample Kolmogorov-Smirnov goodness of fit test was applied by making pair-wise comparisons of the histograms of the aggregated fixation data.[16] Error probabilities of $p < 0.05$ were considered significant. All statistical analyses were carried out with the statistical toolbox of MATLAB 7.1 (Mathworks, Natick, MA).

3. Results

3.1. *Questionnaire data*

Despite the low sample size and novel experimental setting, the analysis of the normalized BCR scores have revealed some interesting trends. Since the goal is to track the learning process of inexperienced subjects, their achievement is fared against that of the more experienced groups. However, even the experts showed some improvement: although Session #1 was relatively easy, nurses as well as experts achieved better scores for more complex movies. The reason might be that subjects were not yet accommodated with the unusual experimental settings. Young residents (presumable with more experience with computers), on the other hand, did manage to handle the novelty of the situation and scored quite well.

Our highlighting method did not facilitate learning. For the naive and the resident groups, there was no significant difference between the highlighted and non-highlighted conditions, while for the nurses and experts, highlighting resulted in significantly *less* scores (H0: scores are equal in Session #3. *P*-values: 0.24, 0.17, 0.04, 0.03, naïve subjects, residents, nurses and experts, respectively). For this reason, in the following analysis, results of only the non-highlighted cases were used.

Fig. 2 depicts the normalized BCR scores for all groups and sessions. The group average responses to the individual video clips (Fig. 2A) showed correlations indicating strong impact of the type of the given clips on the performance independent of expertise level. One extreme case was sample #4 in Session #1 for which the scores of all groups were the highest. Inspection of the separate TP and TN scores revealed that this seizure was much harder to interpret than its initial complexity rate suggested. There were only a few symptoms present and their recognition was difficult. In turn, the imbalance between TP and TN distorted the BCR rate. Since its complexity did not fit the other clips within the same session, we excluded this sample from further analysis.

Fig. 2B depicts the BCR scores averaged over the subjects and the samples within a session (averaging over the video clips might decrease the effect of the peculiarities of the individual video clips). Experts' mean performance was 17–23% lower than the reference's score and was 3–5% higher than naïve observers' scores. The increasing mean scores as well as the decreasing deviations demonstrated the overall improvement for each group (score achieved at session #1 is less than at session #3, two-sample one-sided *t*-test with unequal variance: naïve subjects: $P < 0.0030$, residents: $P < 0.0071$, nurses: $P < 0.0004$, experts: $P < 0.0037$). Nurses' overall performance was lower than that of naïve observers.

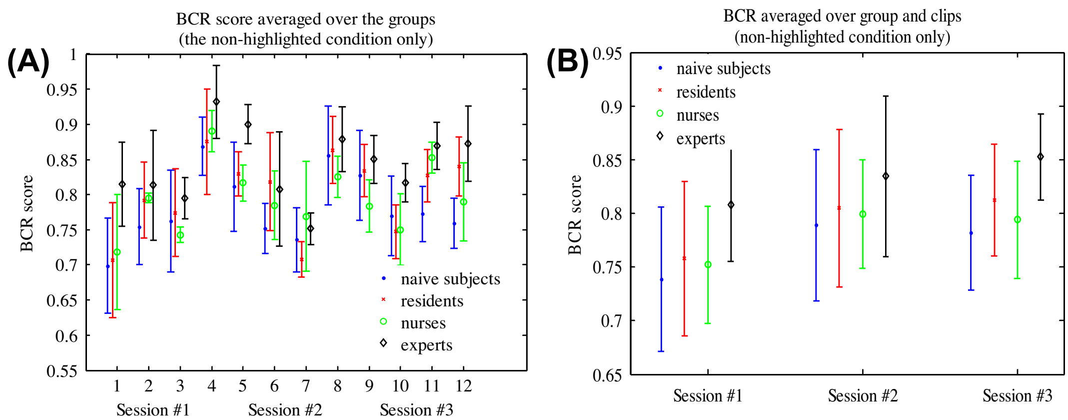

Fig. 2 – (A) Analysis of the questionnaire scores. The BCR scores (mean ± std) averaged over the subjects. There were 4 video clips shown in each session. The correlation of the scores of different groups shows the strong impact of the quality/type of the given clips on the performance. It can also be seen here that each group demonstrates some improvement over the sessions. (B) The BCR scores (mean ± std) averaged over the subjects and video samples within a session (video #4 of session #1 was excluded). Averaging over the video clips may decrease the effect of the peculiarities of the individual video clips. The increasing mean scores as well as the decreasing deviations demonstrate the overall improvement for each group.

3.2. *Analysis of fixation duration distributions*

Since highlighting did not prove to be useful in the learning process, we used the eye-tracking data of only the unmodified (non-highlighted) video samples for further analysis. Due to the large individual differences as well as the diversity of the video clips (differing quality, length, complexity) we pooled the fixation data of subjects belonging to the same group. In this way a distinct fixation duration histogram could be constructed for each group and each session. Histograms of the dwell time (Fig. 3A) indicated that fixations were modulated by expertise. Fixation density values of <200 ms fixation duration were higher for more experienced observers (experts and nurses), whereas inexperienced observers (naïve people and residents) had mostly longer (200–400 ms) dwell times. This difference, however, decreased from session to session. Fig. 3B shows quantification of differences between the fixation distributions. Although all groups moved away from the original fixation pattern during the subsequent sessions, only nurses and experts showed a significant deviation from their initial fixation patterns. Nurses also had a significant evolution toward the experts' patterns.

4. Discussion

The main message of our study is that combining verbalized and visual information is a powerful tool to train subjects to provide better reports on seizure events. Questionnaires seem to serve two purposes. First they provide a contextual frame of reference and help verbalize the observed events. Second, their fast evaluation guarantees the effectiveness of the training method. Besides the analysis of written responses, the analysis of behavioral responses has also revealed some interesting aspects of the learning process. In accord with other studies we have also shown that eye motion characteristics also correlate well with the level of expertise[6,7] or skills.[15,17] Our study is the first using this method in clinical epileptology and the results may lead to novel methods for online evaluation in medical education. The relatively large difference between experts' and reference's performance reflects the subjectivity of seizure description. What is surprising though is that average performance of naïve observers was very close to that of the experts. As typical seizure reports from relatives and lay eye-witnesses are usually unreliable, our findings suggest that reliability of seizure description can be greatly improved by questionnaires containing simple explanation of seizure elements. Interestingly, all groups showed some improvement in the performance and within group variance was decreasing through the three sessions in all groups. It may indicate the decreasing effect of the individual differences on the performance due to the learning process. Another possible factor is that more difficult seizure samples tend to show more concurrent symptoms, so the answers to the different questions were not exclusive.

Experience on seizure recognition cannot be easily verbalized, therefore we were interested if it is possible to guide and facilitate perception of the visual scene by highlighting the most relevant regions of the displayed video clips. However; instead of facilitating the learning process, the highlighted video samples made the observation even more difficult and decreased observers' performance in all groups. The reason for this failure is probably due to the complexity of the task and the particular choice of highlighting. A certain seizure has different parallel ictal elements,[8] so highlighting only one important region might mislead seizure observers. Further studies are needed to see how, under what conditions, and up to what degree can image manipulation be used for improving seizure observation. Let us remark that guiding attention by manipulating visual stimuli has great potentials in visualization, advertisement or education and better results could likely be achieved by alternative highlighting methods.[18] There are other examples using successful highlighting methods in medical diagnostics. Lichtfield et al. assessed the accuracy of chest X-ray reports by "leading"

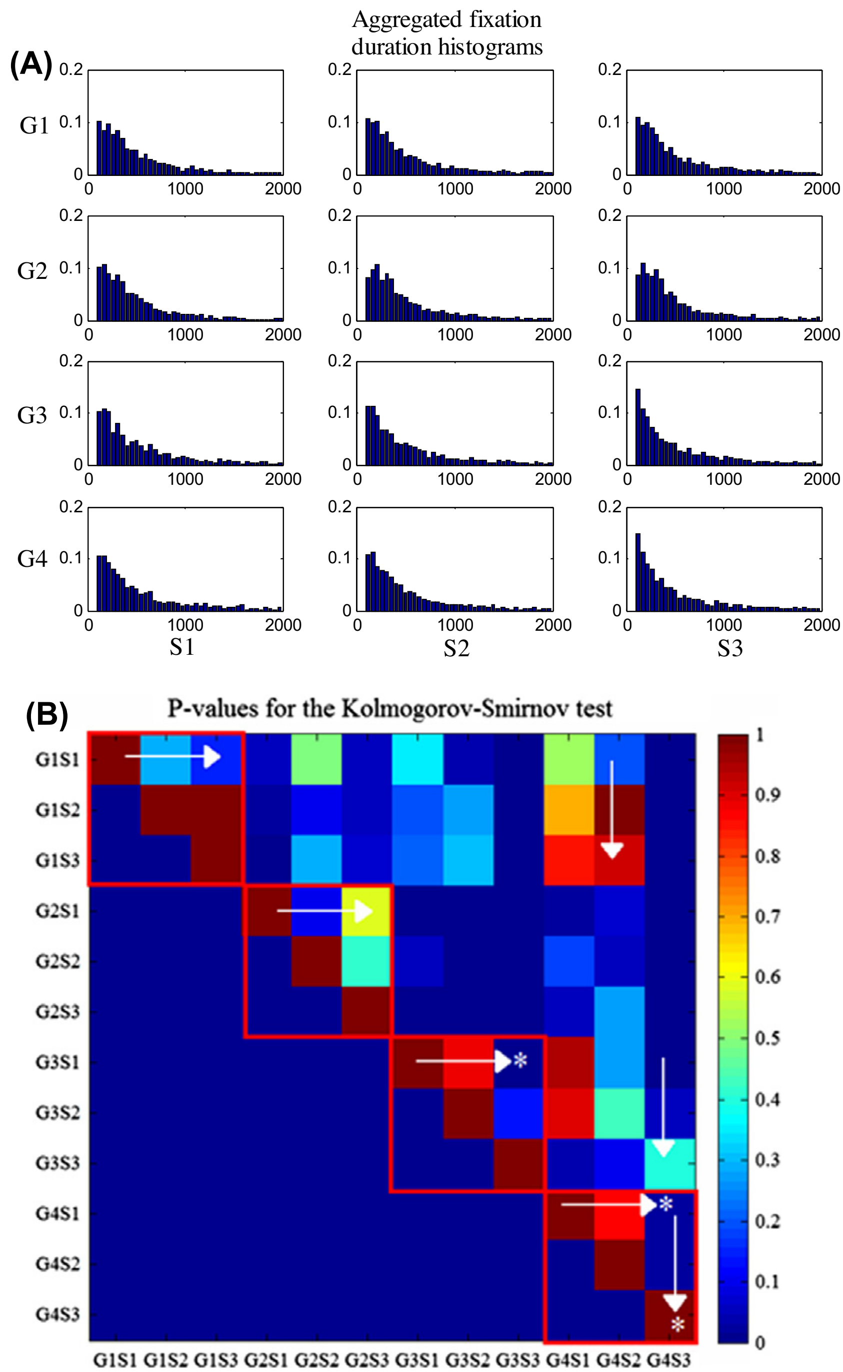

Fig. 3 – Analysis of the fixation duration distributions. (A) Fixation durations of all subjects within a group were aggregated for each session. The distribution of these aggregated durations was approximated by a histogram with 40 bins (one bin is about 50 ms). Even visual inspection may show differences between the shapes of the distributions, e.g. more experiences seem to result in steeper exponential-like shape. (B) To quantify the differences between the histograms a two-sample Kolmogorov-Smirnov goodness of fit test was applied (H0 is that data come from the same distribution). The matrix depicts the corresponding P-values. Horizontal arrows denote the within group improvement: smaller P values mean larger deviation from the initial distribution. Stars denote significant changes at 0.05 level. Vertical arrows in the right-most column indicate that the shapes of the histograms become more similar to the ones of the experts': larger P values mean larger similarity. Red boxes denote session by session differences within each group. Notation: G1 (Group #1), G2, G3 and G4 denotes naïve subjects, residents, nurses and experts, respectively. S1 (Session #1), S2 and S3 denote the subsequent sessions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the residents' eyes by a more experienced radiologist's eye movement and found that this method helped finding more chest nodules.[19] However, this study highlighted stable images and not moving video clips.

Dwell time histograms indicated that fixations were modulated by expertise. Fixation time values were shorter in more experienced observers (experts and nurses), whereas inexperienced observers (naïve people and residents) had mostly longer dwelling times. This difference, however, diminished from session to session reflecting the effect of the learning process. We think that shorter fixation times give more chance to observe more details of a seizure resulting in more precise observation and description. It is important to note, however, that several other factors (e.g., different complexity of seizures) might influence eye-motion characteristics.

Naive observers and residents showed lower performance viewing more complex video samples while the more experienced groups showed improvement even in the last session. Since nurses and experts were not assumed to gain additional knowledge from watching the sample clips, their improvement likely reflects that they still adapted to the unusual conditions of the experimental setup (e.g., answering questionnaires and watching video samples). On the other hand, the lower than expected performance of nurses in the beginning of the experiment could be a result of emotional effects. They personally knew patients shown on the video records and this emotional involvement might affect their attention thus resulting in lower scores. This factor is well known from our everyday experience: parents or relatives of patients with epilepsy frequently have difficulties on objective seizure description. Yet, as even the emotionally attached subjects demonstrated improvement during the short sessions, we believe our concept could be used to train not only health care professionals, but also parents and care givers with no previous experience on seizures.

5. Limitation

Due to the small sample size and the large variability in the complexity of the video samples, the statistical analysis of this pilot study should be taken with care. However, its promising results make it worth to repeat the experiment on larger groups involving parents and caregivers of epileptic patients, as well. For a larger sample size, it would also be useful to study the impact of age on the results using factor analysis.

6. Conclusion

Our pilot study showed that well-established educational questionnaires can train seizure observation of observers with different experience on epilepsy. Improvement of naïve observers' scores and eye-tracking characteristics suggest that this method could be adapted for training of parents, relatives and other caregivers of patients with epilepsy in future. In accord with the primary goal of our study, we plan on creating a novel tutoring program for parents as well as for professionals in education. The program must be easily accessible, fast and efficient in order to gain broad support. In addition, a more specialized program could be introduced in the training of health care professionals. Its practical relevance would be to get more precise and informative reports on epileptic seizures helping neurologists to find the most appropriate diagnosis and choose the best treatment for a given type of epilepsy.

Conflict of interest statement

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines. None of the authors has any conflict of interest to disclose.

Acknowledgement

This work was supported by grant EU Commission PERCEPT (043261) (A.L, G.Sz, and G.Z.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.npbr.2011.10.002.

REFERENCES

1. Noachtar S. Video analysis for defining the symptomatogenic zone. In: Rosenow F, Lüders HO, editors. *Presurgical assessment of the epilepsies with clinical neurophysiology and functional imaging*. New York: Elsevier; 2004. p. 187–200.
2. Fogarasi A, Janszky J, Siegler Z, Tuxhorn I. Ictal smile lateralizes to the right hemisphere in childhood epilepsy. *Epilepsia* 2005;**46**:449–51.
3. Chanlin L. The effects of verbal elaboration and visual elaboration on student learning. *Int J Instructional Media* 1997;**24**:333–9.
4. McKay E. An investigation of text-based instructional materials enhanced with graphics. *Educ Psychol* 1999;**19**:323–35.
5. Scheiter K, van Gog T. Using eye tracking in applied research to study and stimulate the processing of information from multi-representational sources. *Appl Cognit Psychol* 2009;**23**:1209–14.
6. Kocak E, Ober J, Berme N, Melvin WS. Eye motion parameters correlate with level of experience in video-assisted surgery: objective testing of three tasks. *J Laparoendosc Adv Surg Tech* 2005;**15**:575–80.
7. Liu Y, Hsueh PY, Lai J, Sangin M, Nüssli MA, Dillenbourg P. Who is the expert? Analyzing gaze data to predict expertise level in collaborative applications. In: *Proceedings of the 2009 IEEE international conference on Multimedia and Expo, ICME'09 Piscataway*. NJ, USA: IEEE Press; 2009. p. 898–901.
8. Lüders H, Acharya J, Baumgartner C, Benbadis S, Bleasel A, Burgess R, et al. Semiological seizure classification. *Epilepsia* 1998;**39**:1006–13.
9. Fogarasi A, Tuxhorn I, Janszky J, Janszky I, Rasonyi G, Kelemen A, Halasz P. Age-dependent seizure semiology in temporal lobe epilepsy. *Epilepsia* 2007;**48**:1697–702.
10. Fogarasi A, Janszky J, Tuxhorn I. Peri-ictal lateralizing signs in children: blinded multiobserver study of 100 children $\leqslant$12 years. *Neurology* 2006;**66**:271–4.

11. Rásonyi G, Fogarasi A, Kelemen A, Janszky J, Halász P. Lateralizing value of postictal automatisms in temporal lobe epilepsy. *Epilepy Res* 2006;**70**:239–43.
12. Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. *Inform Process Manage* 2009;**45**:427–37.
13. Salvucci DD, Goldberg JH. Identifying fixations and saccades in eye-tracking protocols. In: *Proceedings of the eye tracking research and applications symposium*. New York: ACM Press; 2000. p. 71–8.
14. Pannasch S, Helmert JR, Roth K, Herbold AK, Walter H. Visual fixation durations and saccadic amplitudes: shifting relationship in a variety of conditions. *J Eye Movement Res* 2008;**4**:1–19.
15. Klin A, Jones W, Schultz R, Volkmar F, Cohen D. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch Gen Psychiatry* 2002;**59**:809–16.
16. Chakravarti IM, Laha RG, Roy J. *Handbook of methods of applied statistics*. New York: John Wiley and Sons; 1967. pp. 392–4.
17. Boraston Z, Blakemore SJ. The application of eye-tracking technology in the study of autism. *J Physiol.* 2007;**581**:893–8.
18. Kosara R. Semantic Depth of Field – Using Blur for Focus + Context Visualization. PhD thesis, Vienna University of Technology, Austria; 2001.
19. Litchfield D, Ball LJ, Donovan T, Manning DJ, Crawford T. Viewing another person's eye movements improves identification of pulmonary nodules in chest x-ray inspection. *J Exp Psychol Appl* 2010;**16**:251–62.