american physiological society®
ADVANCES IN PHYSIOLOGY EDUCATION

Adv Physiol Educ 50: 671–678, 2026.
First published April 23, 2026; doi:10.1152/advan.00288.2025

EDUCATION RESEARCH

Prior knowledge of normal ECGs enhances accuracy in ECG interpretation among medical students

Jennifer Edwards,[1] Alistair Gallagher,[1] and W. Joseph MacInnes[2]
[1]School of Medicine, Swansea University, Swansea, United Kingdom and [2]Department of Computer Science, Swansea University, Swansea, United Kingdom

Abstract

Electrocardiograms (ECGs) are frequently misdiagnosed, with potentially serious consequences, yet evidence for effective teaching strategies remains limited. Eye-tracking offers insight into the perceptual processes behind accurate interpretation. We conducted an eye-tracking study among medical students to assess whether priming with normal ECGs improves diagnostic accuracy. We hypothesized that participants viewing normal ECGs before interpreting unknown ECGs would perform more accurately. This between-subjects priming study presented students with one of three image types before an ECG task: normal ECGs with instruction that they were normal, normal ECGs without instruction, or chest X-rays (control). The diagnostic task comprised 30 ECGs across six conditions. The study was built with Experiment Builder, and eye movements were recorded with an EyeLink 1000 Plus system. Students performed well above chance, with accuracy comparable to the 54% reported in meta-analysis. Atrial fibrillation was most often identified correctly, whereas left bundle branch block was the most challenging. Across groups, faster responses and finally fixating the primary abnormality (PA) predicted accuracy, although inaccurate participants spent longer overall on the PA area of interest. Students who were informed that the priming ECGs were normal responded faster and more accurately than those primed unknowingly or with chest radiographs. ECG-primed groups also showed longer fixations in nonrelevant areas, whereas explicitly primed students showed no significant related secondary abnormality fixation sequences and more dispersed spatial attention. These early findings suggest that targeted priming with normal ECGs may support learning. Variation in diagnostic difficulty across conditions indicates that tailored training approaches may be required.

NEW & NOTEWORTHY This is the first study to investigate the effect of priming medical students with normal ECGs on the accuracy of subsequent ECG interpretation. Using eye-tracking methodology, we characterized visual behaviors associated with improved diagnostic performance among primed students. Our findings demonstrate that exposure to normal ECG patterns before analyzing an abnormal trace influences accuracy and gaze allocation, suggesting that priming may enhance pattern recognition and inform the design of educational strategies in ECG training.

ECG interpretation; eye-tracking; medical education

INTRODUCTION

Electrocardiograms (ECGs) are one of the most widely used diagnostic medical tests (1), yet they are frequently misinterpreted, with median accuracy of interpretation in educational settings just 54% (2). These errors result in incorrect management, poor outcomes, and avoidable deaths (3). There is a recognized need for better, standardized, ECG training for medical professionals (2, 3); however, evidence is lacking as to which teaching techniques are effective (2–4). This is a consideration for educators across medical imaging domains, including radiology, dermatology, pathology, and cardiology (5). Yet, given the subconscious nature of much of medical decision-making, experts lack insight into successful interpretation strategies that may aid educational approaches (6).

Discovering anomalies in medical imaging is a difficult visual task similar to visual search and foraging (7). Human attention is a key resource in how we build information in service to a task. The human brain cannot process the full amount of perceptual information it receives, so attention allows individuals to prioritize information that is likely to support their current goals and inhibit that which is not. For visual tasks eye movements are an excellent observable proxy for the allocation of attention and are easily recorded with eye trackers (8). Eye-tracking data captured during ECG interpretation can reveal visual search and decision-making strategies to feed into educational approaches or guide interventions.

Experts display an efficient search strategy prioritizing relevant information, and eye-tracking research often describes and compares novice and expert approaches to interpretation (9, 10). Experts are believed to deploy a "top-down" approach where visual cues leading to diagnosis are focused on faster and, importantly, their significance is understood

Correspondence: J. Edwards (jenniferljedwards@gmail.com).
Submitted 25 November 2025 / Revised 27 January 2026 / Accepted 17 April 2026

(9, 10). They are thought to use their prior knowledge of what is normal to quickly identify abnormalities before scrutinizing them further (4). Conversely, novices may frequently employ exhaustive search strategies (11), fixating on abnormalities without understanding their significance, and this is often described as a "bottom-up" approach (10, 12). One study compared groups by accuracy rather than expertise (13) and found that although most participants fixated on salient features, the inaccurate group failed to understand their significance. This suggests that the inaccurate group, regardless of expertise, exhibited a novicelike bottom-up approach, whereas the accurate group used an expertlike top-down approach, potentially reflecting differences in metacognitive awareness and control over diagnostic reasoning (14).

Various additional measures have been proposed to improve ECG interpretation among learners, but none has gained widespread support or is strongly evidence based. The most common recommendation for novices is the use of a systematic, structured approach when interpreting (3, 4). Others also suggest teaching a better understanding of the physiology of depolarization and repolarization driving ECG morphologies, rather than focusing on simply memorizing them (3, 4).

Looking to research into interpreting other medical imaging modalities for guidance, there appears to be merit in placing greater emphasis on normal cases within the curriculum. Fundamentally, normal images constitute the majority of what is encountered in medical practice, yet medical image perception training typically focuses on abnormalities (15). Furthermore, experts are theorized to achieve higher diagnostic accuracy through their knowledge of both normal and abnormal presentations (9, 13), and it has therefore been recommended that novices be exposed to more normal clinical cases to develop robust nonpathological internal schemata (9). Eye-tracking research from the complex visual domain of radiological interpretation shows that expert interpreters compare normal and abnormal regions within symmetrical anatomical structures (16), highlighting the value of an internal reference for normality.

We sought to advance the field by investigating the impact of exposure to normal ECGs by priming medical students with normal ECGs before they diagnosed potentially abnormal ones. Our hypothesis was that interpreters who knowingly viewed normal ECGs before interpreting ECGs with an unknown diagnosis would be more accurate. We further predicted that analysis of eye movement patterns, comparing behaviors between accurate and inaccurate respondents and by priming groups, would show how attentional allocation leads to successful diagnostic decisions.

MATERIALS AND METHODS

Participant Recruitment and Consent

Thirty-four students ($n = 34$) were recruited from years 1–3 of the 4-yr Graduate Entry Medicine program at Swansea University. An undergraduate degree in any subject is a requirement for entry, meaning students have a wide range of academic and often professional experience. Most participants had not studied ECG interpretation before this course, and only one had read ECGs as part of a prior clinical role. All underwent similar instruction under the Swansea University curriculum, although some variance is expected between years. Participants were recruited via word of mouth and were eligible to take part as long as they could read fine detail on a screen a foot away and focus on a screen for an hour in duration without suffering any ill effects.

All participants were briefed on the format of the experiment but not the purpose. All gave written informed consent for participation. Participants were debriefed after the study as to its purpose. Data collected were anonymized, with consent information stored separately from responses to minimize risk in the event of data breach. Data were stored on a password-protected computer file, with hard copies stored in locked filing cabinets. Ethical approval was obtained from the Swansea University Science and Engineering ethics committee, approval number 1 2023 8627 7375.

Experimental Setup and Procedures

This was a between-subject research design with three priming groups comprising a review of medical images. The experiment was built with Experiment Builder, and EyeLink 1000 Plus was the eye tracker used to capture eye movements (17). The EyeLink 1000 Plus is a widely used, high-precision video-based eye-tracking system with a sampling rate of up to 2,000 Hz (samples per second) and manufacturer-reported spatial accuracy of $\sim 0.1^\circ$ of visual angle (dva) (17). To ensure data quality, a 9-point calibration and validation procedure was performed before recording. Calibration was repeated until average spatial error was $<0.5^\circ$ and maximum error was $<1^\circ$. Participants were positioned with a chin rest at a fixed distance of 50 cm to minimize head movement and maintain calibration stability. Drift correction and brief calibration checks were performed between trials to ensure continued accuracy throughout the experiment. Experiment Builder and DataViewer were then used to implement the millisecond-accurate experiment and extract the fixation data for further analysis.

Participants were randomly assigned to one of three priming groups upon arrival: *group 1* "X-Ray" ($n = 12$), *group 2* "Normal" ($n = 11$), and *group 3* "Normal-Told" ($n = 11$). All three groups were shown five medical images before the experiment phase, with five normal ECGs used for the priming phase of *group 2* and *group 3* and five normal chest X-rays shown to participants in control *group 1*. *Groups 2* and *3* differed in that *group 2* was not told that the ECGs being shown were normal, with no abnormalities, whereas *group 3* was told that the ECGs were normal. After this priming phase, participants were able to move away from the eye tracker and take a break before beginning the experimental phase.

In the experimental phase, once participants were recalibrated to the eye-tracking apparatus, instructions for the ECG interpretation task were displayed on screen along with the six possible diagnosis options (outlined below). Participants clicked to progress to the first ECG once ready. They had up to 60 s to review the ECG before automatic progression to the answer screen, although participants were instructed that they could click to progress sooner if they had completed review. Participants then had up to 60 s to select from one of the six answer options, with automatic progression to the next ECG and no answer recorded if they failed to select in time. Thirty ECGs were shown during the experimental phase, with a maximum possible duration of 60 min.

ECG Selection and Coding

Five conditions were chosen from a list of common ECG conditions proposed that medical students be proficient in by graduation (3). Five ECGs from each condition category were included, plus five ECGs displaying no abnormalities. Each participant was therefore asked to evaluate the following conditions in the experiment phase: normal sinus rhythm (NSR), 1st degree heart block (1DHB), atrial fibrillation (AFIB), left bundle branch block (LBBB), inferior ST-elevation myocardial infarction (ISTEMI), and anterolateral ST-elevation myocardial infarction (ALSTEMI). All ECGs were in 12-lead format, displayed in a 3 × 4 grid with a rhythm strip.

All medical images used were anonymous, and as they were sourced from educational and training websites we deemed fair use for our experiment (18–21). Areas of abnormality had been delineated by clinicians, which we used to inform classification of our ECGs (see Table 1 and Supplemental Fig. S1).

For each category of ECG, waveform patterns were determined and classified as one of the following: primary abnormality (PA), related secondary abnormality (RSA), unrelated secondary abnormality (USA), and normal (N). Before the experiment, Experiment Builder was used to create areas of interest (AOIs) around each lead, coded as one of these four categories. Delineating AOIs at the lead level is a common approach used in many eye-tracking studies of ECG interpretation (9, 22, 23), although analyzing eye movements at the waveform level may become feasible as eye-tracking technology continues to develop (24). A separate set of AOIs defined only by lead name were also created. This coding was used to inform our eye-tracking analysis.

Data Capture and Analysis

Eye movements were recorded during both the priming and experimental phases, and we analyzed the number of fixations in AOIs, duration of time spent in each AOI, and area of last fixation before decision. Eye-tracking metrics such as fixation count, fixation duration, and dwell time within AOIs are well-established proxies for visual attention and cognitive processing, with demonstrated validity in medical image interpretation research (25).

We also captured noneye movement data, including the amount of time it took to form a diagnosis, which was denoted "response time," the participant's diagnosis for each ECG, and whether they were correct or incorrect.

Students were also asked the following questions in a post-experiment questionnaire: *1*) Have you done any courses/training in reading ECGs? If so, please specify. *2*) Do you have any practical experience reading ECGs? *3*) Were there any strategies you used to assess the ECGs? They were also invited to write any feedback they had on the study more generally. Thirty-one student questionnaires were collected.

Data were analyzed with mixed-effects models in RStudio (26) with the packages lme4 (27), ANOVA results with the package car (28), pairwise comparisons and error bars for marginal means with estimated marginal means (emmeans) (29), and graphs with ggplot2 (30). The analysis for accuracy used a generalized linear mixed-effect model with a binomial distribution, priming group and diagnosis as independent variables, and participant and image as nested random factors. Response time analysis was similar but used the gamma distribution and included accuracy as an independent variable to account for speed accuracy trade-offs. A second model of accuracy was used for eye-tracking analysis with priming group, average fixation duration, and last fixation before decision as independent variables. This analysis used a reduced dataset with normal ECGs removed since they did not contain an area of PA. All pairwise analysis and plotted values were expected marginal means from the emmeans package to better reflect the results from the above models.

For fixation pattern analysis, fixations were coded according to their spatial location. Analyses for two different codings were conducted: one based on ECG lead regions and the second based on defined AOIs. Each region was assigned a letter code (A, B, C) to simplify fixation sequences, for example a sequence visiting leads I, II, II, and I was represented as "ABBA." This coding enabled the identification of recurring subsequences, or temporal patterns, within and across groups. The GroupStr R package, using the CommonPatt function, was employed to detect subsequences present in at least 30% of trials across six groups, defined by three experimental conditions and diagnostic accuracy, whether participants were correct or incorrect. We also analyzed the fixations for each trial considering the spatial dispersion of

Table 1. *Summary of areas of interest used for coding electrocardiograms in the study*

	Primary Abnormality (PA)	Related Secondary Abnormality (RSA)
1st degree heart block (1DHB)	Abnormality: PR interval > 200 ms Location: All leads	Abnormality: N/A Location: N/A
Atrial fibrillation (AFIB)	Abnormality: Irregularly irregular rhythm, absent P waves Location: All leads	Abnormality: Absence of isoelectric baseline Location: All leads
Left bundle branch block (LBBB)	Abnormality: Broad monophasic R wave, with absence of Q wave Location: Leads I, aVL, V5–V6	Abnormality: QRS > 120 ms Location: All leads
Inferior ST-elevation myocardial infarction (ISTEMI)	Abnormality: ST segment elevation Location: II, III, aVF	Abnormality: ST segment depression Location: aVL
Anterolateral ST-elevation myocardial infarction (ALSTEMI)	Abnormality: ST segment elevation Location: I, aVL, V1–V6	Abnormality: ST segment depression Location: II, III, aVF

For each electrocardiogram (ECG) condition, the primary abnormality (PA) and related secondary abnormality (RSA) are shown, along with their locations on the standard 12-lead ECG (18). N/A, not applicable.

the fixations. These measures often provide insight into cognitive strategies; for example, different tasks often show different dispersion patterns even with the same image (E. Akinola and W. J. MacInnes, unpublished observations). Participants' data were tested with bivariate contour ellipse area (BCEA), which measures the area of the ellipse needed to contain 68% of the fixations for a given trial. This provides a quantitative assessment of gaze variability and has been used as indicator of performance and attentional distribution (31).

Alpha was set at $P = 0.05$ for all tests.

RESULTS

Analysis of all student interpretations in the experimental phase (normal and abnormal ECGs) showed that priming group was significant in predicting accuracy ($\chi^2 = 10.7$, df = 2, P = 0.005). The main effect of diagnosis was also significant ($\chi^2 = 49.3$, df = 5, $P < 0.001$), but the interaction did not significantly improve the model ($\chi^2 = 11.5$, df = 10, P = 0.323). Paired comparison of priming groups showed that the main factor driving this was the difference between *group 2* and *group 3* ($z = -3.3$, $P = 0.004$). *Group 2* had the worst performance of all three groups (47.6% accurate), compared to 56.4% accuracy in *group 1* and 61.1% accuracy in *group 3* (Fig. 1). As such, we were able to reject our null hypothesis and based on these early results surmise that priming with normal ECGs does lead interpreters to analyze ECGs more accurately.

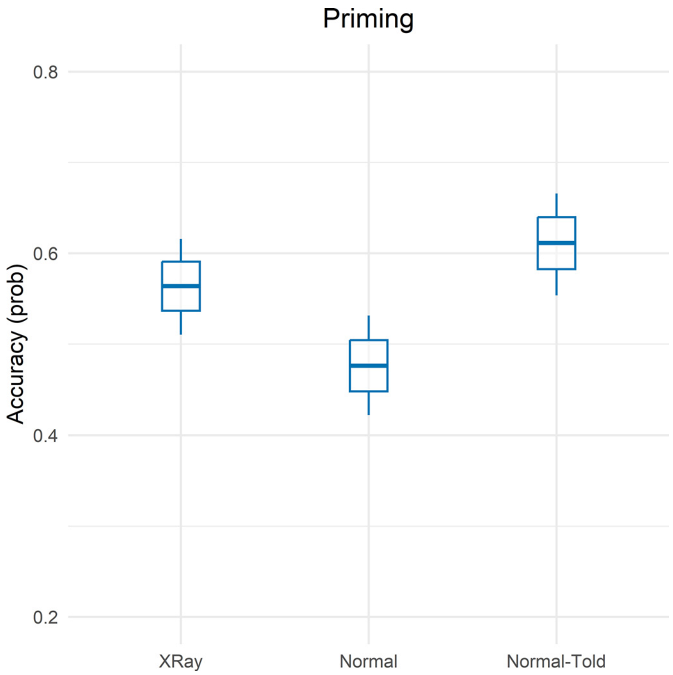

Figure 1. Boxplot showing accuracy of electrocardiogram (ECG) interpretation (percentage of correct responses during the experimental phase) by priming group. Participants were assigned to X-Ray (n = 12), Normal (n = 11), or Normal-Told (n = 11) groups. During the priming phase, the X-Ray group viewed 5 X-ray images. The Normal group viewed 5 ECGs without being informed that they were normal before interpretation, whereas the Normal-Told group viewed 5 ECGs and were informed that they were normal. Plotted means are the expected marginal means from the mixed-effect models. The box limits represent the SE, and the whiskers represent 95% confidence interval (CI).

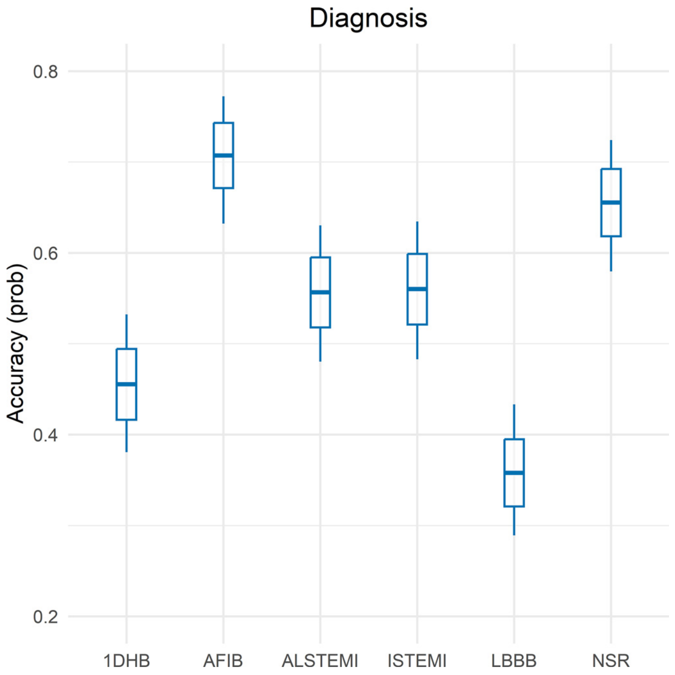

Figure 2. Boxplot showing accuracy of electrocardiogram (ECG) interpretation (percentage of correct responses during the experimental phase) by diagnosis across all participants (n = 34). Diagnoses were 1st degree heart block (1DHB), atrial fibrillation (AFIB), anterolateral ST-elevation myocardial infarction (ALSTEMI), inferior ST-elevation myocardial infarction (ISTEMI), left bundle branch block (LBBB), and normal sinus rhythm (NSR). Plotted means are the expected marginal means from the mixed-effect models. The box limits represent the SE, and the whiskers represent 95% confidence interval (CI).

For the main effect of diagnosis, we observed that participants were most likely to diagnose AFIB correctly (70.7%), followed by NSR (65.5%). Most participants were able to correctly diagnose ALSTEMI (55.6%) and ISTEMI (56.0%). LBBB was the most challenging to diagnose (35.8%), followed by 1DHB (45.5%) (Fig. 2). Pairwise comparisons showed that most diagnoses were performed better than LBBB, including NSR ($z = -3.614$, $P = 0.004$), AFIB ($z = 6.198$, $P < 0.001$), ALSTEMI ($z = 3.618$, $P = 0.004$), ISTEMI ($z = 3.662$, $P = 0.003$), and NSR ($z = -5.337$, $P < 0.001$), with the lone nonsignificant exception being 1DHB ($z = 1.802$, $P = 0.464$). 1DHB was in addition significantly worse than both AFIB ($z = -4.532$, $P < 0.001$) and NSR ($z = -3.614$, $P = 0.004$). No other pairs differed significantly.

We also examined response times and included both priming group and accuracy as independent variables, with the latter to test for potential speed accuracy trade-offs in the groups. Accuracy was found to predict faster response times ($\chi^2 = 26.7$, df = 1, $P < 0.001$), but there was no effect of priming group ($\chi^2 = 3.0$, df = 2, $P = 0.225$) nor did the interaction quite reach significance ($\chi^2 = 5.8$, df = 2, $P = 0.055$). We have chosen to include the accuracy for the separate groups here despite the lack of significant interaction, simply for descriptive clarity. In *group 1* there was a response time difference for inaccurate versus accurate ($z = -4.6$, $P < 0.001$), with the inaccurate group slower by 10 s. *Group 3* also showed a response time difference for inaccurate versus accurate ($z = -3.1$, $P = 0.002$), with the inaccurate group slower by 5 s. In *group 2* there was found to be no significant difference ($z = -1.2$, $P = 0.248$) in response time, though the direction of the difference was the same as the other groups, with inaccurate trials being slower.

To determine whether there was a learning effect or difference in learning between the priming groups, we analyzed accuracy by trial order and priming groups as independent variables. Other variables of the mixed model were left unchanged. Because stimuli were presented in random order for each participant, trials were divided into five equal bins of 20% each, creating an "order" variable from 1 to 5 indicating the trial's position in the experiment. The priming group effect remained significant ($\chi^2 = 10.7$, df = 2, $P = 0.005$), consistent with the effect observed previously (see Fig. 1). The main effect of accuracy also increased significantly over the course of the experiment ($\chi^2 = 10.4$, df = 4, $P = 0.034$), with the probability of a correct response averaged across all participants rising from 49.5%, 50.7%, 54.0%, 55.7%, to 63.9% across the five quintiles. These probabilities reflect overall accuracy across all priming groups. The priming group × order interaction was not significant ($\chi^2 = 4.9$, df = 8, $P = 0.767$), suggesting that the advantage for *group 3* did not change significantly as the experiment progressed.

Temporal patterns of fixations between AOIs were evaluated by finding common subsequences of regions visited. Subsequences that were common in at least 30% of trials were compared across priming conditions and trial accuracy. Across all groups, common fixation sequences tended to be longer in inaccurate trials than in accurate ones, particularly within area PA, which was the dominant region of repeated fixations. In *group 1*, inaccurate participants exhibited longer fixation sequences in both PA and N, whereas accurate participants showed shorter, more targeted viewing sequence patterns. *Group 2* displayed a similar pattern for PA but the opposite trend for N, with accurate trials showing longer sequences there. In *group 3*, accurate participants again demonstrated extended fixations in N and showed no significant or sustained fixation sequences in RSA, whereas inaccurate trials concentrated for longer sequences in both PA and RSA. Overall, inaccurate trials were characterized by longer, less selective fixation sequences across regions. The same analysis was applied to sequences of fixations based on raw lead labels, but the only common sequences meeting threshold were short-duration sequences on the rhythm strip.

We examined fixation patterns for each trial with respect to their spatial dispersion. The mean BCEA values for *groups 1* (603,585) and *2* (604,861) did not significantly differ ($t < 1.0$). However, *group 3* exhibited a significantly greater mean BCEA (647,209) compared with both *group 2* ($t = -4.4$, $P < 0.001$) and *group 1* ($t = -3.9$, $P < 0.001$).

We also carried out nonsequential or spatial analysis of fixations comparing different AOIs, focused on priming group, the last place participants looked before making a diagnosis, and average fixation duration and their impact on accuracy. We only analyzed nonnormal ECGs from our experimental phase ($n = 25$), as normal ECGs only feature "normal" coded AOIs. We observed a main effect of AOI last fixation type before decision ($\chi^2 = 12.4$, df = 3, $P = 0.006$). All other main effects and interactions were not found to be significant in this analysis.

Carrying out pair comparisons showed that those who look in PA as last fixation are most likely to be accurate, as determined by the main effect. When comparing those looking in PA AOIs as last fixation compared to an area of RSA, PA was found to be significantly more predictive of correctness than RSA ($P = 0.009$, $z = -3.2$). In comparison with N AOIs, the ability of PA to predict accuracy approached, but did not reach, significance ($z = 0.23$, $P = 0.098$). PA as final area of fixation versus USA was found to not be significant ($z = -3.154$, $P = 0.290$) (Fig. 3). This is demonstrated by two exemplar heatmaps, showing a greater concentration of fixations in PA by accurate participants compared to inaccurate participants in the final 5 s (Fig. 4 and Fig. 5).

DISCUSSION

Overall, our student participants were able to perform the task at well above chance across conditions and not dissimilar from the 54% accuracy in diagnosis found in meta-analysis (2). AFIB was most frequently diagnosed correctly, whereas LBBB was the most challenging. Three behaviors were identified among accurate students, and these were largely independent of priming group: faster response times predicted greater accuracy; accurate students were more likely to view an area of PA immediately before making a diagnosis; and finally, accurate students tended to have overall shorter fixation sequences when inspecting the PA AOI.

Priming group did impact a number of our results: for example, we found that students who were explicitly informed they were viewing normal ECGs (*group 3*)

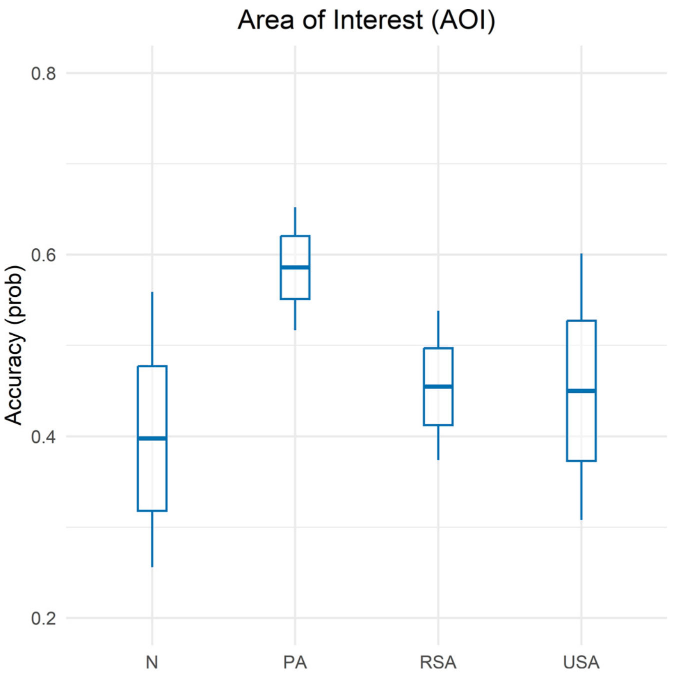

Figure 3. Boxplot showing accuracy of electrocardiogram (ECG) interpretation (percentage of correct responses during the experimental phase) by region of final visual fixation, as determined by eye-tracking, across all participants ($n = 34$). Regions of interest were normal (N), primary abnormality (PA), related secondary abnormality (RSA), and unrelated secondary abnormality (USA). Plotted means are the expected marginal means from the mixed-effect models. The box limits represent the SE, and the whiskers represent 95% confidence interval (CI).

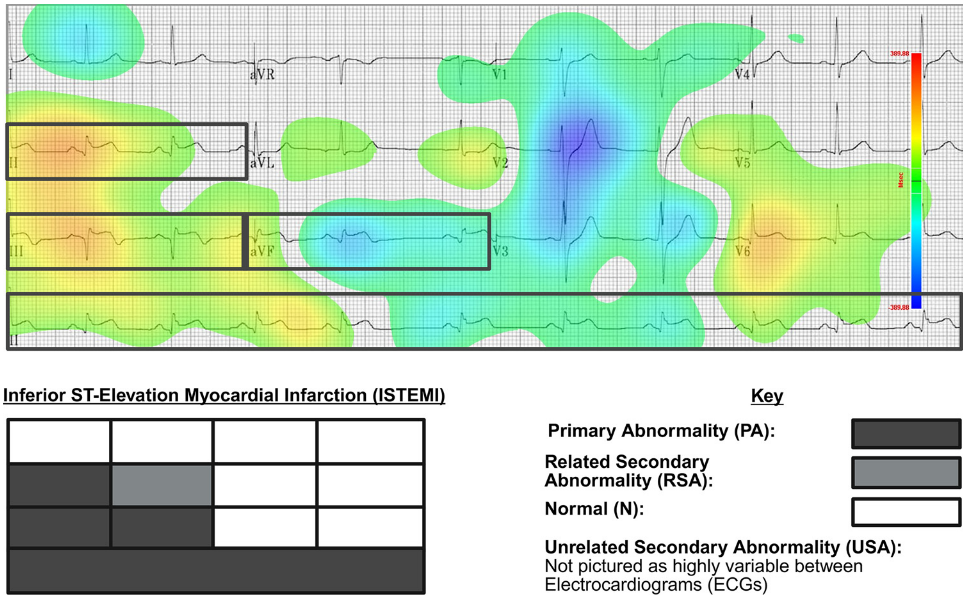

Figure 4. Heat map of eye movements during the final 5 s of a trial showing an inferior ST-elevation myocardial infarction (ISTEMI) electrocardiogram (ECG), across all participants ($n = 34$). Warmer colors (red/yellow) indicate visual fixation by participants who correctly interpreted the ECG, primarily within areas of interest (AOIs) corresponding to the primary abnormality (PA) (leads II, III, and aVF). Cooler colors (blue/green) indicate fixation by participants with incorrect interpretations, distributed outside these regions. A schematic is shown alongside to illustrate AOI coding for this ECG diagnosis, and PA AOIs are marked on the ECG. Figure created with a licensed version of BioRender.com.

demonstrated faster and more accurate interpretations than those who viewed chest X-rays or normal ECGs unknowingly. Accuracy also increased over the course of the experiment, rising from 49.5% in early trials to 63.9% in later trials, independent of priming group. This suggests a learning or familiarization effect rather than a fatigue effect, as performance improved rather than declined across trials. The advantage of group did not change over time, as the priming × order interaction was not significant. We note, however, that the randomized stimuli in our experiment design might not allow detection of more subtle interactions between trial order and priming. Groups primed with ECGs (*2* and *3*) tended to look longer at nonrelevant areas during trials in which they answered correctly than during incorrect trials. Notably, *group 3* showed no detectable RSA fixation sequences at the 30% threshold. Finally, spatial fixation analysis further indicated that the attention of *group 3* was more widely dispersed across trials.

Our findings suggest that familiarization with normal ECG morphologies may be beneficial in medical education, with normal priming leading to both greater accuracy and distinct visual search behaviors among accurate primed students as determined by temporal and fixation analyses. Both groups showed a tendency to spend longer sequences in nonrelevant AOIs during accurate trials, including normal ECGs, potentially reflecting efforts to confirm their choice or conclusively eliminate alternatives. Notably, only those knowingly exposed to normal ECGs did not demonstrate significant sequences within RSA, suggesting a more focused comparison between

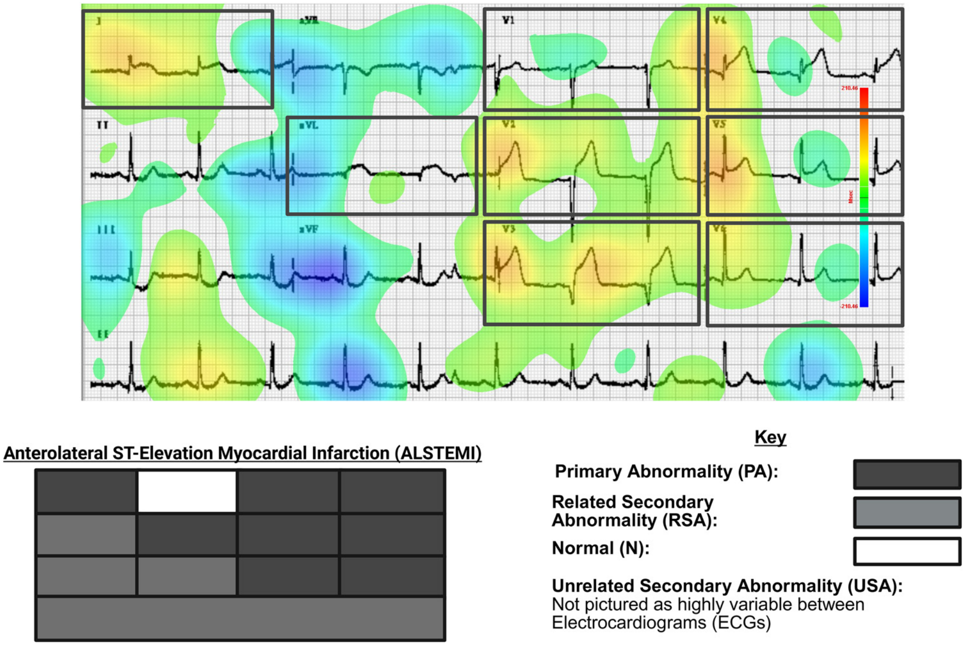

Figure 5. Heat map of eye movements during the final 5 s of a trial showing an anterolateral ST-elevation myocardial infarction (ALSTEMI) electrocardiogram (ECG), across all participants ($n = 34$). Warmer colors (red/yellow) indicate visual fixation by participants who correctly interpreted the ECG, primarily within areas of interest (AOIs) corresponding to the primary abnormality (PA) (leads I, aVL, and V2–V5). Cooler colors (blue/green) indicate fixation by participants with incorrect interpretations, distributed outside these regions. A schematic is shown alongside to illustrate AOI coding for this ECG diagnosis, and PA AOIs are marked on the ECG. Figure created with a licensed version of BioRender.com.

the PA and N areas. Additionally, spatial fixation analysis further indicated that this group exhibited a wider dispersion of attention across trials, a pattern that may reflect an expertlike, top-down scanning strategy to obtain an overall impression before focusing on areas of diagnostic interest.

There are several caveats and areas for future investigation. The effect of known versus unknown exposure to normal ECGs must be further explored, given that overall the unaware priming group were the least accurate and that different search strategies were identified in our two experimental groups. Other impacts of exposure to normal morphologies should also be investigated. A radiology study investigating the effects of varying proportions of normal images without eye-tracking found that diagnostic performance depended on image distribution: conditions with 30% normal cases achieved higher sensitivity, whereas those with 70% normal cases achieved higher specificity, indicating a sensitivity-specificity trade-off that would need to be accounted for when including greater exposure to normal images in curriculums (15).

A further study employing a visual diagnosis task reported that students who compared pathological images with normal examples demonstrated superior diagnostic accuracy for focal diseases (such as those with lesions limited to 1 site) compared to those without such comparisons. In contrast, no significant group difference was observed for diffuse diseases (those showing pathology bilaterally) (32). The results suggest that referencing normal images helps learners extract diagnostically relevant information in focal conditions only. More research is needed to understand why this effect occurs in radiology; however, there is exciting potential to explore similar mechanisms in ECG interpretation, particularly given the requirement to diagnose cardiovascular conditions with both focal and diffuse presentation on an ECG. We would also suggest that future research should explore whether priming with normal ECG waveforms immediately before interpretation has an impact on accuracy over the longer term.

We also extended existing ECG eye-tracking research by comparing accurate and inaccurate novice behaviors, rather than focusing on expertise alone. Although better understanding experts' interpretative approach is acknowledged to be important (9), research has yet to fully explain the underlying mechanisms and patterns of expert visual behavior. Proficiency is likely at least partially attributable to sheer experience (13). However, we identified three expertlike, top-down behaviors among accurate novices, independent of priming group. They were more likely to reach their diagnosis quickly, a trait considered reflective of experts' advanced search and diagnostic strategies (9). They spent less time inspecting PA AOI when accurate, in contrast to inaccurate participants spending longer inspecting PA AOI, perhaps reflecting uncertainty and inattentional blindness of key features (33). They were also more likely to focus on an area of PA in the 5 s before diagnosis, suggesting an understanding of the significance of the area. These findings align with previous research comparing eye movement behaviors based on accuracy among experts and novices. As in our study, those with higher accuracy exhibited top-down processing behaviors typical of experts, regardless of expertise level (13). Overall, our results show that novices' lack of expertise does not preclude them from exhibiting expertlike performance. Future research analyzing successful and unsuccessful interpretative behaviors in larger, more diverse novice groups could reveal accessible strategies to expedite learner competence.

Finally, in line with previous research, our findings show variation in diagnostic difficulty across cardiac conditions. Although varying participant bases and conditions prevent direct comparisons, we can still identify key trends arising. LBBB was mostly misdiagnosed by student-only groups while correctly diagnosed by a mixed-experience cohort, whereas AFIB and NSR were among the most accurately diagnosed across studies (12, 13). STEMIs were consistently diagnosed with moderate accuracy (12, 13). Although we did not analyze eye-tracking behaviors between conditions, scan path analysis in other research suggests that interpretation strategies can vary (13). Future research could focus on these differences to inform targeted teaching approaches by condition.

Our study had some strengths and limitations. Eye-tracking studies in general benefit from a larger sample size to provide more statistical power. However, given the specialized nature of our study among medical students, and the challenges associated with recruitment, we are nonetheless pleased to have secured one of the larger sample sizes of students in an eye-tracking study of ECG interpretation. Like much research in this field, our participants were limited to medical students. Future work could include a wider range of health care professionals who read ECGs in role, such as nurses, paramedics and physician associates, and draw comparisons between and among these groups. We utilized a forced-choice format, where participants chose one of the six possible diagnoses. This enabled discrete data collection and robust statistical analysis, while maintaining immersion in the eye-tracking apparatus. However, choosing from a predetermined list may have been less challenging than originating an independent response, potentially resulting in inflated rates of correct diagnoses. Additionally, the multichoice format influenced some interpretation strategies, with several students reporting adjusting their approaches based on the diagnoses they were expecting. Future research could validate findings against a free-answer format. Limitations also stem from the controlled research setting. Participants noted that using a computer screen restricted some of their usual interpretative techniques, such as measuring R-R intervals with paper. Unlike real clinical settings, our environment lacked distractions and external pressures, potentially limiting the generalizability of findings. Use of portable eye trackers in future studies could record eye movements during ECG interpretation in less controlled contexts, such as a clinical or classroom setting, to assess differences in behavior and ensure that strategies to improve accuracy are practically applicable.

To conclude, our early results suggest that exposure to normal morphologies may have beneficial impact on interpretation accuracy and identify unique visual search behaviors among those primed. We also uncovered new insights into how novices interpret ECGs. That faster diagnosis and ascribing of significance to key abnormalities, behaviors typical of experts, were observed in novices shows the potential to exhibit expertlike performance despite a lack of experience. We also observed variations in diagnostic difficulty across conditions, with certain conditions, such as LBBB, proving more challenging to diagnose. These results suggest that future educational strategies could be tailored to specific conditions to improve diagnostic accuracy. We believe our work paves the way for further research into how exposure to normal cases could boost ECG interpretation skills among trainee health care professionals, which could also be applied across fields including radiology, dermatology, and pathology.

DATA AVAILABILITY

The experimental data supporting the findings of this study are publicly available on the Open Science Framework (OSF) at the following link: https://doi.org/10.17605/OSF.IO/WPX7N.

SUPPLEMENTAL MATERIAL

Supplemental Fig. S1: https://doi.org/10.17605/OSF.IO/WPX7N.

ACKNOWLEDGMENTS

Figures 4 and 5 were created with a licensed version of BioRender.com.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

J.E., A.G., and W.J.M. conceived and designed research; J.E. and A.G. performed experiments; J.E. and W.J.M. analyzed data; J.E. and W.J.M. interpreted results of experiments; W.J.M. prepared figures; J.E. and W.J.M. drafted manuscript; J.E. and W.J.M. edited and revised manuscript; J.E., A.G., and W.J.M. approved final version of manuscript.

REFERENCES

1. **Baranchuk A**, **Chiale PA**, **Green M**, **Caldwell JC.** Surface electrocardiogram remains alive in the XXI century. *World J Cardiol* 6: 34–35, 2014. doi:10.2174/1573403X1003140522160529.
2. **Cook DA**, **Oh SY**, **Pusic MV.** Accuracy of physicians' electrocardiogram interpretations: a systematic review and meta-analysis. *JAMA Intern Med* 180: 1461–1471, 2020. doi:10.1001/jamainternmed.2020.3989.
3. **Antiperovitch P**, **Zareba W**, **Steinberg JS**, **Bacharova L**, **Tereshchenko LG**, **Farre J**, **Nikus K**, **Ikeda T**, **Baranchuk A.** Proposed in-training electrocardiogram interpretation competencies for undergraduate and postgraduate trainees. *J Hosp Med* 13: 185–193, 2018. doi:10.12788/jhm.2876.
4. **Breen CJ**, **Kelly GP**, **Kernohan WG.** ECG interpretation skill acquisition: a review of learning, teaching and assessment. *J Electrocardiol* 73: 125–128, 2022. doi:10.1016/j.jelectrocard.2019.03.010.
5. **Kok EM**, **van Geel K**, **van Merriënboer JJ**, **Robben SG.** What we do and do not know about teaching medical image interpretation. *Front Psychol* 8: 309, 2017. doi:10.3389/fpsyg.2017.00309.
6. **Al-Moteri MO**, **Symmons M**, **Plummer V**, **Cooper S.** Eye tracking to investigate cue processing in medical decision-making: a scoping review. *Comput Hum Behav* 66: 52–66, 2017. doi:10.1016/j.chb.2016.09.022.
7. **Wolfe JM.** Visual search: how do we find what we are looking for? *Annu Rev Vis Sci* 6: 539–562, 2020. doi:10.1146/annurev-vision-091718-015048.
8. **Wolfe JM**, **Horowitz TS.** Five factors that guide attention in visual search. *Nat Hum Behav* 1: 0058, 2017. doi:10.1038/s41562-017-0058.
9. **Wood G**, **Batt J**, **Appelboam A**, **Harris A**, **Wilson MR.** Exploring the impact of expertise, clinical history, and visual search on electrocardiogram interpretation. *Med Decis Making* 34: 75–83, 2014. doi:10.1177/0272989X13492016.
10. **Broadbent M**, **Horsley M**, **Birks M**, **Persaud N.** Comparing novice and expert nurses in analysing electrocardiographs (ECGs) containing critical diagnostic information: an eye tracking study of the development of complex nursing visual cognitive skills. In: *Current Trends in Eye Tracking Research*, edited by Horsley M, Eliot M, Knight B, Reilly R. Springer, 2014, p. 297–316.
11. **Davies A**, **Brown G**, **Vigo M**, **Harper S**, **Horseman L**, **Splendiani B**, **Hill E**, **Jay C.** Exploring the relationship between eye movements and electrocardiogram interpretation accuracy. *Sci Rep* 6: 38227, 2016. doi:10.1038/srep38227.
12. **Sqalli MT**, **Al-Thani D**, **Elshazly MB**, **Al-Hijji M.** Interpretation of a 12-lead electrocardiogram by medical students: quantitative eye-tracking approach. *JMIR Med Educ* 7: e26675, 2021. doi:10.2196/26675.
13. **Davies A.** Examining Expertise through Eye Movements: a Study of Clinicians Interpreting Electrocardiograms (PhD thesis). University of Manchester, 2018.
14. **Croskerry P.** A universal model of diagnostic reasoning. *Acad Med* 84: 1022–1028, 2009. doi:10.1097/ACM.0b013e3181ace703.
15. **van Geel K**, **Kok EM**, **Aldekhayel AD**, **Robben SG**, **van Merriënboer JJ.** Chest X-ray evaluation training: impact of normal and abnormal image ratio and instructional sequence. *Med Educ* 53: 153–164, 2019. doi:10.1111/medu.13756.
16. **Anderson B**, **Shyu CR.** A preliminary study to understand tacit knowledge and visual routines of medical experts through gaze tracking. *AMIA Annu Symp Proc* 2010: 21–25, 2010.
17. **SR Research Ltd.** EyeLink 1000 Plus [Binocular tower] (Online). SR Research Ltd, 2013. https://www.sr-research.com/eyelink-1000-plus/ [2025 Mar 27].
18. **ECG Library.** Life in the Fast Lane (Online). https://litfl.com/ecg-library/ [2025 Mar 27].
19. **ECGpedia.** ECGpedia: The online ECG library (Online). https://en.ecgpedia.org/ [2025 Mar 27].
20. **HQmeded.** HQmeded: Medical education resources (Online). https://hqmeded.com/ [2025 Mar 27].
21. **Radiopaedia.** Radiopaedia: The free radiology resource (Online). https://radiopaedia.org/ [2025 Mar 27].
22. **Bond R**, **Finlay DD**, **Breen C**, **Boyd K**, **Nugent CD**, **Black ND**, **Macfarlane PW**, **Guldenring D.** Eye tracking in the assessment of electrocardiogram interpretation techniques. *Comput Cardiol* 39: 581–584, 2012.
23. **Breen CJ**, **Bond R**, **Finlay D.** An evaluation of eye tracking technology in the assessment of 12 lead electrocardiography interpretation. *J Electrocardiol* 47: 922–929, 2014. doi:10.1016/j.jelectrocard.2014.08.008.
24. **Davies A**, **Mueller J**, **Horseman L**, **Splendiani B**, **Hill E**, **Vigo M**, **Harper S**, **Jay C.** How do healthcare practitioners read electrocardiograms? A dual-process model of electrocardiogram interpretation. *Br J Card Nurs* 14: 1–19, 2019. doi:10.12968/bjca.2019.0073.
25. **Wu CC**, **Wolfe JM.** Eye movements in medical image perception: a selective review of past, present and future. *Vision (Basel)* 3: 32, 2019. doi:10.3390/vision3020032.
26. **Posit Team.** RStudio: Integrated Development Environment for R (Online). Posit Software, PBC, 2025. http://www.posit.co/ [2026 Mar 24].
27. **Bates D**, **Mächler M**, **Bolker B**, **Walker S.** Fitting linear mixed-effects models using lme4. *J Stat Soft* 67: 1–48, 2015. doi:10.18637/jss.v067.i01.
28. **Fox J**, **Weisberg S.** *An R Companion to Applied Regression* (3rd ed.). Sage, 2019.
29. **Lenth R**, **Piaskowski J.** emmeans: Estimated marginal means, aka least-squares means (Online). 2025. https://rvlenth.github.io/emmeans/.
30. **Wickham H.** *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, 2016.
31. **Ramey MM**, **Henderson JM**, **Yonelinas AP.** The spatial distribution of attention predicts familiarity strength during encoding and retrieval. *J Exp Psychol Gen* 149: 2046–2062, 2020. doi:10.1037/xge0000758.
32. **Kok EM**, **de Bruin AB**, **Robben SG**, **van Merriënboer JJ.** Learning radiological appearances of diseases: does comparison help? *Learn Instr* 23: 90–97, 2013. doi:10.1016/j.learninstruc.2012.07.004.
33. **Chetverikov A**, **Kuvaldina M**, **MacInnes WJ**, **Jóhannesson ÓI**, **Kristjánsson Á.** Implicit processing during change blindness revealed with mouse-contingent and gaze-contingent displays. *Atten Percept Psychophys* 80: 844–859, 2018. doi:10.3758/s13414-017-1468-5.