*Corresponding Author's Email: roger@mail.tpa.edu.tw
Proceedings of the World Conference on Teaching, Learning, and Education
Vol. 3 Issue. 1, 2026, pp. 01- 12
DOI: https://doi.org/10.33422/worldtle.v3i1.1827

ISSN: 3030-1580 online

Diamond Scientific Publishing

An Eye-Tracking Study of Visual Strategies in Crime Scene Investigations

Liang-Yi Chung[1], and Rong-Chi Chang2[*]

[1]Chihlee University of Technology
[2]Taiwan Police College

Abstract

This study examines how prior knowledge and formal instruction shape visual attention during crime scene investigation (CSI). Using a quasi-experimental pretest-posttest design, we compared a prior-knowledge cadet group (n = 21), a novice cadet group (n = 30), and an expert reference group of active forensic investigators (n = 13). Participants viewed authentic crime-scene photographs while gaze data were recorded for key-evidence and distractor areas of interest (AOIs). Eleven eye-tracking metrics and Lag Sequential Analysis (LSA) were used to characterize attention allocation and gaze transitions. Results show that prior knowledge guided evidence-oriented viewing before instruction, that both student groups improved after one semester of CSI training through different visual-learning pathways, and that eye-tracking detected differences not captured by written tests. The findings support differentiated CSI instruction and the use of formative gaze feedback in law-enforcement education.

Keywords: crime scene investigation, eye tracking, visual attention, expert–novice differences

1. Introduction

Crime scene investigation (CSI) is a cognitively demanding task in which officers must rapidly separate probative evidence from visually salient but irrelevant information (Fisher & Fisher, 2012). Forensic viewing is therefore not passive seeing; it is hypothesis-guided visual search shaped by professional schemas (Chang & Tsai, 2022). When attention is captured by irrelevant cues, confirmation bias or tunnel vision can influence decisions across the investigation process (Fahsing & Ask, 2016; Rossmo, 2009).

Eye-tracking studies have demonstrated substantial differences in visual search strategies between experts and novices across professional domains, including radiology (Kundel & Nodine, 1975), aviation (Bellenkes et al., 1997), and airport security screening (McCarley et al., 2004). Experts tend to allocate attention more efficiently to task-relevant features, whereas novices are more likely to be guided by bottom-up visual salience (Gegenfurtner et al., 2011). Although visual attention is central to CSI practice, empirical eye-tracking research in forensic

training remains limited. Recent work has therefore called for systematic studies that examine how training shapes gaze behavior in applied forensic contexts (Stainer et al., 2023).

Taiwan Police College offers a natural comparison between two student populations: Special Examination cadets, who often enter with social, legal, or investigative experience, and Regular Diploma cadets, who are typically recent high-school graduates without prior investigative training. Comparing these groups before and after a semester-long CSI course makes it possible to examine both the initial role of prior knowledge and the learning effects of formal instruction.

This study addresses two research questions:

RQ1: Do differences in prior knowledge lead to significant differences in visual attention allocation when viewing crime-scene photographs?

RQ2: Does formal CSI instruction change gaze behavior differently across learners with different background knowledge?

To align these questions with the analytic design, RQ1 was examined through pretest group comparisons of AOI-based gaze metrics between the prior-knowledge and novice groups. RQ2 was examined through pretest-posttest comparisons within each student group. Lag Sequential Analysis (LSA) was used to characterize gaze-transition strategies among key evidence, distractor, and outside-AOI regions. Written test scores were included as a conventional assessment benchmark, allowing gaze-based indicators of perceptual strategy to be compared with explicit knowledge performance.

By linking prior knowledge, instructional change, gaze-transition strategy, and written assessment, this study aims to clarify how forensic visual expertise develops and how eye-tracking evidence may inform differentiated, evidence-based CSI training.

2. Literature review

2.1 Visual Search Based on Knowledge

In 1967, Yarbus' pioneering study showed that how individuals scan the same images is affected by their goals. In later work this idea has been generalized to the concept of knowledge-driven visual search, where top-down attentional guidance to task-relevant features is enabled by domain schemas acquired through training and experience, over and above bottom-up salience signals (Nodine & Mello-Thoms, 2010). Meta-analyzing 30 eye-tracking experiments from professional domains, Gegenfurtner et al. (2011) found that knowledge reliably predicts more efficient and targeted fixation patterns, especially in complex diagnostic imaging tasks.

2.2 Differences in Visual Performance Between Experts and Novices

Differences between expert and novice visual search have been studied best in medical imaging. Kundel and Nodine (1975) showed that expert radiologists could reach accurate interpretations of chest radiographs within fractions of a second through a "holistic global perception" before systematic search, while novices relied on sequential, exhaustive scanning. Comparable asymmetries have been found in sports (Abernethy & Russell, 1987), medical education (Wilson et al., 2011), and forensic document examination (Mnookin, 2008). The basic cognitive mechanism is schema-guided feature expectation (Kosslyn & Koenig, 1992): experts pre-activate templates that amplify target salience and diminish irrelevant distractors.

2.3 Learning by Perception and Making Skills Automatic

Theories of skill acquisition suggest that as information is proceduralized through deliberate practice, cognitive resource requirements are reduced and performance becomes more

automatic (Anderson, 1983; Gibson, 1969). Fitts and Posner (1967) identify a continuum from cognitive (conscious control) to autonomous (automatic) stages. Accuracy was constant, while automatization in visual search was characterized by a decrease in number and duration of fixations. This pattern occurs only in expert performance domains where individuals engage in deliberate practice (Ericsson, 2006). Dreyfus and Dreyfus (1986) proposed a five-level competence continuum (novice → advanced beginner → competent → proficient → expert), which provides a theoretical basis for interpreting differences in gaze patterns between groups of participants.

3. Method

3.1 Participants

A total of 64 participants completed both test sessions and met the data-quality criterion of at least 80% valid gaze coverage per image: a prior-knowledge group of special examination cadets (n = 21; M age = 30.2 years, SD = 4.7), a novice group of regular diploma cadets (n = 30; M age = 19.8 years, SD = 1.2), and an expert reference group of active forensic investigators (n = 13; M age = 38.6 years, SD = 6.3; M experience = 12.4 years). A priori and post-hoc G*Power analyses indicated adequate power for the primary group comparisons and paired pretest-posttest analyses. All participants had corrected-to-normal vision, no color deficiency, and provided written informed consent. The study was approved by the National Chengchi University Research Ethics Committee (Approval No. NCCU-REC-202505E050).

Table 1. Participant Characteristics

Group	n	Mean Age (SD)	Gender (M/F)	Background
Prior-knowledge	21	30.2 (4.7)	18 / 3	Adult professionals; prior legal/investigative knowledge
Novice	30	19.8 (1.2)	22 / 8	Recent high school graduates; no investigative background
Expert (reference)	13	38.6 (6.3)	11 / 2	Active forensic officers; mean 12.4 yrs field experience

3.2 Stimuli and Areas of Interest (AOI)

A panel of three senior forensic officers, each with more than 20 years of field experience, selected 20 authentic color photographs representing five crime-scene categories: indoor arson, gunshot/windshield damage, homicide, suspected narcotics manufacturing, and physical assault. The photographs were divided into two equivalent sets of 10 images for the pretest and posttest. Pilot testing indicated no significant difference in set difficulty, $t(22) = 0.83$, $p = .415$.

For each image, two AOI zones were defined according to forensic relevance. Zone_1 represented key evidence areas, such as V-pattern burn marks, bullet entry or exit points, blood spatter, and suspected contraband. Zone_2 represented background or distractor areas, such as furnishings, perimeter tape, and architectural elements. The AOI boundaries were independently reviewed by the three expert raters. Inter-rater reliability, assessed using the Jaccard Similarity Coefficient for area overlap, showed high agreement, M = 0.89, SD = 0.04.

3.3 Apparatus and Procedure

Gaze data were recorded using a Tobii 4C eye tracker with a 90 Hz sampling rate. Stimuli were displayed at 1920 × 1080 resolution on a 24-inch Full HD monitor at an approximate viewing distance of 60 cm. A nine-point calibration procedure was administered before each session, and participants who failed to meet calibration criteria were excluded.

Student participants completed the pretest in the first week of the semester and the posttest in the penultimate week, following approximately 18 weeks of instruction in the Crime Scene Processing and Evidence Collection course. The expert reference group completed the same eye-tracking task once as a benchmark and was not included in the longitudinal pretest-posttest comparison. In each session, participants viewed 10 crime-scene photographs at their own pace and then answered two to three multiple-choice questions assessing evidence recognition, forensic cue interpretation, and basic procedural judgment. Participants were not informed of the predefined AOI locations.

Before analysis, blinks, off-screen gaze points, and tracker-loss periods were removed. No interpolation was applied to missing gaze segments. Image trials with less than 80% valid gaze coverage were excluded from gaze-metric and LSA analyses. Valid fixations were mapped onto Zone_1, Zone_2, or outside-AOI areas. For LSA, fixation sequences were converted into categorical transition codes, and invalid or missing samples were not coded as transitions.

3.4 Statistical Analysis

Eleven AOI-based eye-tracking metrics were calculated for each participant and image: Total Time Spent (TTS), Total Fixation Duration (TFD), Total Fixation Count (TFC), Duration before First Fixation Arrival (FFA), Average Fixation Duration (AFD), Percent Time Spent (PTS), Proportion of Fixation Duration (PFD), Percentage of Fixation Count (PFC), First Pass Time (FPT), Second Pass Time (SPT), and Revisit Fixation Duration (RFD).

Because several AOI metrics were conceptually related, the main results were interpreted at the metric-cluster level rather than as isolated outcomes. TTS, TFC, PTS, FFA/FPT, and SPT/RFD were emphasized as primary indicators of attention volume, search effort, relative allocation, early orienting, and revisitation, respectively. The remaining metrics were used as supporting evidence and are reported in the supplementary material.

Group comparisons at pretest were examined using independent-samples *t*-tests, and pretest-posttest changes within each student group were examined using paired-samples *t*-tests. Response accuracy was analyzed using chi-square tests. All tests were two-tailed with $\alpha = .05$. To reduce Type I error across multiple AOI metrics, Holm-Bonferroni corrections were applied. Effect sizes are reported as Cohen's d with 95% confidence intervals.

Lag Sequential Analysis (LSA) was conducted using GSEQ 5.1 (Bakeman & Quera, 2011) to examine transitions among Zone_1, Zone_2, and outside-AOI regions. Yule's Q was used as the association strength index, with $z > 1.96$, $p < .05$, indicating significant transitions. Analyses were conducted with and without self-loops to distinguish within-zone revisiting from between-zone shifts.

Given the modest subgroup sample sizes and the image-specific nature of several effects, image-level findings were interpreted as theoretically meaningful indicators rather than definitive population-level estimates. Future analyses should extend the present t-test, chi-square, and LSA approach by specifying mixed-effects models with participants and images as random effects, and group, testing phase, AOI zone, and their interactions as fixed effects. Such models would more directly account for image-level heterogeneity and help distinguish generalizable gaze-pattern differences from effects driven by particular scene characteristics.

4. Results

To improve interpretability, the eye-tracking results are reported using theoretically meaningful metric clusters rather than as isolated individual measures. As shown in Table 2, the eleven AOI metrics were organized into four gaze-behavior dimensions: attention volume, relative allocation, search timing, and revisitation or sustained processing. Lag Sequential Analysis

(LSA) was reported separately as an index of gaze-transition strategy. This structure allows the results to show whether different metrics converge on the same visual-attention pattern while reducing overinterpretation of any single image-level effect.

Table 2. Metric Clusters Used to Interpret Eye-Tracking Results

Metric cluster	Metrics included	Gaze-behavior meaning	Role in the Results section
Attention volume	TTS, TFD, TFC	Amount of visual attention allocated to an AOI	Indicates whether participants devoted more gaze time or fixation activity to evidence or distractor regions
Relative allocation	PTS, PFD, PFC	Proportion of total gaze devoted to an AOI	Shows whether evidence or distractor regions received priority after accounting for overall viewing differences
Search timing	FFA, FPT	Speed and early-stage orienting to an AOI	Captures how quickly participants detected or first processed key evidence
Revisitation / sustained processing	SPT, RFD	Return visits and extended inspection	Indicates repeated checking, uncertainty, exhaustive search, or sustained evidence evaluation
Transition strategy	LSA transitions, Yule's Q	Sequential movement among Zone_1, Zone_2, and outside-AOI regions	Describes whether gaze patterns reflected unstable switching, evidence-oriented search, or flexible scene-adaptive exploration

4.1 Test Score Performance

Both groups improved comparably across the instructional period. Pretest scores did not differ significantly (prior-knowledge: $M = 49.63$, $SD = 11.81$; novice: $M = 44.96$, $SD = 18.00$; $t(49) = -0.88$, $p = .383$, $d = -0.30$). Posttest scores similarly showed no significant group difference (prior-knowledge: $M = 58.89$; novice: $M = 54.16$; $t(49) = -0.86$, $p = .394$, $d = -0.29$). Both groups gained approximately 9 points over the semester. This score equivalence provides a critical interpretive context: observable differences in gaze behavior between groups cannot be attributed to differential knowledge acquisition as measured by conventional testing.

4.2 Pretest Differences Between Groups

For RQ1, pretest group comparisons examined whether prior knowledge shaped visual attention allocation before formal instruction. Across the attention-volume and relative-allocation clusters, the prior-knowledge group showed a more evidence-oriented viewing pattern than the novice group. This pattern was illustrated by the representative indoor arson image, where the prior-knowledge group allocated more gaze to Zone_1, with higher TTS, TFD, TFC, and PTS values, whereas the novice group spent proportionally more time on Zone_2. In contrast, search-timing indicators such as FFA did not show a reliable group difference, suggesting that the groups did not differ primarily in initial evidence detection but in sustained attention and revisitation of evidence regions. To avoid overemphasizing single-image effects, the main text reports the convergent pattern across metric clusters, with complete image-level statistics provided in the supplementary material.

This pattern indicates that prior knowledge influenced how participants prioritized and maintained attention on forensic evidence, rather than simply how quickly they first located

relevant regions. Detailed image-level statistics, including means, confidence intervals, and effect sizes, are provided in Supplementary Table S1.

4.3 Posttest Group Differences

After instruction, group differences in Zone_1 attention persisted and in some cases increased. This posttest pattern was illustrated by representative image-level evidence from the attack-scene image, where effect sizes for TTS, PTS, PFD, and FPT exceeded the corresponding pretest range. This pattern suggests that identical instruction did not eliminate the initial advantage of prior knowledge; instead, the advantage became more visible in gaze behavior.

4.4 Pretest-Posttest Changes in the Prior-Knowledge Group

For the prior-knowledge group, pretest-posttest changes were most evident in the attention-volume and search-timing clusters, suggesting increased efficiency rather than reduced evidence priority. Absolute Zone_1 gaze metrics declined after instruction, including TTS, TFC, FFA, and AFD, while proportional evidence attention remained relatively stable. This pattern suggests perceptual automatization: learners required fewer fixation resources to maintain evidence-oriented attention. Complete image-level statistics for the prior-knowledge group are provided in Supplementary Table S2.

4.5 Pretest-Posttest Changes in the Novice Group

For RQ2, pretest-posttest comparisons in the novice group examined whether formal instruction changed learners' gaze allocation and search strategy. The learning pattern was most clearly illustrated by the bullet-hole glass image, where reductions were observed across attention-volume, relative-allocation, and revisitation metrics. Specifically, posttest decreases in TTS, PTS, and SPT suggest that novices no longer relied on prolonged or repeated scanning of Zone_1 after instruction. Instead, their viewing pattern became more selective and efficient, consistent with a transition from exhaustive search toward targeted evidence inspection.

The pattern was not uniform across all scene types. In the arson scene, pretest-posttest changes were weaker, indicating that visual learning may depend on the clarity of diagnostic cues and the spatial reasoning demands of the evidence type. Detailed image-level statistics for the novice group are reported in Supplementary Table S3.

4.6 Lag Sequential Analysis

LSA examined how participants shifted attention between evidence and distractor regions across images, groups, and testing phases. Yule's Q and $z > 1.96$ were used to identify significant transitions; full image-level matrices are provided in the supplementary material.

At pretest, the novice group showed frequent bidirectional Zone_1↔Zone_2 transitions, indicating unstable switching between evidence and distractor regions. This pattern suggests an exhaustive search mode in which learners had not yet developed stable forensic criteria for prioritizing evidence. The prior-knowledge group showed fewer cross-zone transitions and a stronger tendency toward evidence-oriented scanning, suggesting an emerging schema-guided search strategy. The expert reference group showed fewer significant sequential links, indicating flexible, scene-adaptive exploration rather than rigid two-zone scanning.

At posttest, the prior-knowledge group moved closer to the expert pattern, with fewer significant cross-zone transitions and less repetitive evidence–distractor switching. The novice group shifted from unsystematic bidirectional switching toward more directional Zone_2→Zone_1 transitions, suggesting that learners began to use contextual background information to orient toward evidence. However, their strategy remained more rule-based and less flexible than that of the prior-knowledge and expert groups.

The LSA results indicate a developmental continuum in gaze strategy: novices initially showed unstable evidence-distractor alternation, prior-knowledge learners showed more evidence-oriented search, and experts displayed flexible scene-adaptive exploration. These patterns were strongest in scenes with visually distinct diagnostic cues and more variable in complex scenes such as arson, where spatial inference is required.

4.7 Hierarchical Summary of Gaze-Behavior Findings

To shift the emphasis from individual scenes to convergent gaze-behavior patterns, Table 3 summarizes the findings by research focus, analysis level, metric-cluster evidence, and interpretation. This alignment clarifies how specific metrics converged within broader gaze-behavior dimensions and how those dimensions support the interpretation of each research question. This hierarchical presentation highlights patterns that were supported by multiple gaze-behavior indicators while preserving the detailed image-level statistics as representative evidence.

Table 3. Hierarchical Summary of Gaze-Behavior Findings

Research focus	**Analysis level**	**Convergent evidence by metric cluster**	**Interpretation**
RQ1: Prior knowledge	Pretest group comparison	**Attention volume**: Zone_1 TTS, TFD, and TFC favored prior-knowledge cadets; **relative allocation**: Zone_1 PTS favored prior-knowledge cadets, whereas Zone_2 PTS favored novices	Prior knowledge supported evidence-oriented viewing before instruction
RQ2: Prior-knowledge learning	Pretest-posttest	**Attention volume/search timing**: absolute Zone_1 metrics such as TTS, TFC, FFA, and AFD declined; **relative allocation**: proportional evidence attention remained stable	Training promoted perceptual automatization
RQ2: Novice learning	Pretest-posttest	**Attention volume**: TTS declined; **relative allocation**: PTS declined; **revisitation/sustained processing**: SPT declined after instruction	Novices shifted from exhaustive to targeted search
Posttest group gap	Posttest group comparison	**Relative allocation/first-pass processing**: posttest effects in TTS, PTS, PFD, and FPT exceeded the pretest effect-size range	Identical instruction did not fully close the initial visual-skill gap
LSA strategy	Sequential analysis	**Transition strategy**: LSA/Yule's Q patterns showed fewer cross-zone transitions in trained/prior-knowledge patterns and flexible exploration in experts	Visual search developed from unstable switching toward schema-guided exploration

Table 3 indicates that the central findings converged across complementary levels of evidence: AOI-based metrics showed where attention was allocated, whereas LSA clarified how attention moved across evidence and distractor regions. Together, these patterns support the interpretation that CSI visual expertise develops through both evidence-oriented allocation and more organized gaze-transition strategies.

5. Discussion

5.1 Prior Knowledge as a Visual Attention Architect

The pretest data show that prior knowledge structures visual attention before formal CSI instruction. The prior-knowledge group's higher Zone_1 fixation metrics (d = 0.63-0.71) suggest an emerging crime-scene cue-recognition schema that helps differentiate task-relevant from task-irrelevant information (Nodine & Mello-Thoms, 2010). In contrast, novices were more vulnerable to bottom-up distractor capture in Zone_2 (d = 0.72).

This gaze advantage appeared without a corresponding advantage on written test scores: both student groups performed similarly on multiple-choice questions before and after instruction. The dissociation indicates that conventional tests capture explicit declarative knowledge but may miss implicit perceptual-strategic competence. Eye tracking therefore offers a diagnostic window into tacit investigative skill.

5.2 Qualitatively Distinct Learning Pathways

Identical instruction produced different gaze changes in the two student groups. For the prior-knowledge group, training appeared to support automatization: absolute fixation volume decreased while proportional attention to evidence was maintained, consistent with perceptual learning accounts (Gibson, 1969; Goldstone, 1998). Learners with richer schemas may integrate procedural rules more efficiently, reducing the cognitive load of each visual decision.

For the novice group, training produced strategic reorganization. Before instruction, elevated TTS, SPT, and RFD reflected exhaustive scanning without clear stopping criteria. After instruction, novices searched more selectively because explicit visual decision rules helped them decide when a forensic feature had been adequately inspected. This pattern aligns with Berbaum et al.'s (1990) account of stopping criteria in visual search.

These dual pathways imply that one instructional approach will not serve all learners equally. Beginners need concrete visual criteria and repeated guided practice. Learners with stronger prior knowledge need complex, time-pressured, and collaborative cases that help them refine and automate existing schemas.

5.3 The Matthew Effect and Its Training Implications

The increase in gaze disparities from pretest (|d| = 0.63-0.71) to posttest (|d| = 0.72-0.94) is consistent with a Matthew Effect in forensic visual skill learning. If a uniform curriculum benefits prepared learners more strongly than novices, standard instruction may unintentionally preserve or widen initial skill gaps.

At the same time, the novice group showed meaningful absolute improvement, with the clearest representative evidence appearing in ballistic images (d = 1.05-1.23). The issue is therefore not whether novices learn, but whether they receive enough structured support to prevent early differences from becoming durable performance gaps. Case-based practice, explicit visual-rule instruction, and formative gaze feedback may help reduce this risk, a point consistent with Russ and Safford's (2022) pilot work on forensic eye-tracking feedback.

5.4 Implementation Roadmap for Gaze-Feedback Training

We propose integrating gaze feedback into CSI training as a formative, process-oriented tool. Aggregated heatmaps and scanpath replays can serve as perceptual mirrors, allowing students to compare their search behavior with expert benchmarks. Metrics such as duration before first fixation arrival can support rapid evidence detection, while percent time spent can help instructors address distractor resistance. Sessions should be low-stakes, repeated periodically, and adapted to learner background: novices need explicit rule-based visual templates, whereas knowledge-rich learners need complex, time-pressured scenarios that promote automatization.

Responsible implementation also requires institutional safeguards. Trainees should receive clear consent information and assurance that gaze data will be used for formative learning rather than personnel evaluation. Raw scanpaths should be retained only for the training cycle, while anonymized aggregate data may support program-level benchmarking. Expert benchmarks should be audited for demographic and crime-type representativeness, and instructors should be calibrated to interpret gaze metrics consistently rather than reducing competence to heatmap appearance.

5.5 Limitations and Future Inquiry

Several limitations should be acknowledged. Although the sample size was adequate for detecting medium-to-large effects in the primary comparisons, it remains modest for estimating subgroup differences at the individual-image level. The single-institution sample also limits generalizability, and age was partially confounded with prior social, legal, and professional experience.

The 90 Hz screen-based eye tracker was appropriate for fixation-level analysis but limited finer-grained oculomotor measures such as microsaccades, saccade dynamics, and pupil-linked cognitive load. In addition, static two-dimensional photographs provided experimental control but could not fully represent embodied crime-scene investigation involving movement, evidence handling, communication, and time pressure.

Future studies should use larger multi-institutional samples, higher-resolution or wearable eye trackers, no-intervention control groups, delayed posttests, and VR-based crime-scene simulations. Analytically, future work should apply mixed-effects models with participants and images as random effects to better account for image-level variance and distinguish generalizable training effects from scene-specific effects.

6. Conclusion

This study provides multi-dimensional eye-tracking evidence that prior knowledge, formal instruction, and their interaction shape visual attention in CSI training. Three conclusions follow: prior knowledge guides forensic gaze before formal training; instruction produces different learning pathways in knowledge-rich and novice learners; and conventional written tests can miss perceptual-strategic differences that eye tracking reveals. These findings support differentiated CSI instruction and evidence-type-specific training design.

Future work should examine whether these gaze-based training effects generalize to larger, multi-institutional, and more realistic training environments. In particular, VR-based simulations, wearable eye tracking, delayed posttests, and mixed-effects modeling could help evaluate the ecological validity, causal impact, and long-term retention of perceptual-skill training in CSI contexts.

Overall, the study argues for a shift in forensic education from uniform knowledge delivery toward differentiated perceptual skill development. Used responsibly, eye tracking can become

a scalable diagnostic and training tool for improving the accuracy, efficiency, and fairness of crime-scene investigation practice.

Acknowledgment

This work was supported by the National Science and Technology Council, Taiwan. (grant number NSTC 114-2410-H-261-0011).

References

Abernethy, B., & Russell, D. G. (1987). Expert-novice differences in an applied selective attention task. *Journal of Sport & Exercise Psychology, 9*(4), 326–345. https://doi.org/10.1123/jsp.9.4.326

Anderson, J. R. (1983). *The architecture of cognition*. Harvard University Press.

Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. Cambridge University Press. https://doi.org/10.1017/CBO9781139017343

Bellenkes, A. H., Wickens, C. D., & Kramer, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine, 68*(7), 569–579.

Berbaum, K. S., Franken, E. A., Jr., et al. (1990). Satisfaction of search in diagnostic radiology. *Investigative Radiology, 25*(2), 133–140. https://doi.org/10.1097/00004424-199002000-00006

Chang, R.-C., & Tsai, M.-J. (2022). Visual behavior patterns of successful decision makers in crime scene photo investigation: An eye tracking analysis. *Journal of Forensic Sciences*. https://doi.org/10.1111/1556-4029.14970

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine*. Simon and Schuster.

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683–703). Cambridge University Press. https://doi.org/10.1017/CBO9780511816796.038

Fahsing, I. A., & Ask, K. (2016). The making of an expert detective. *Psychology, Crime & Law, 22*(3), 203–223. https://doi.org/10.1080/1068316X.2015.1077249

Fisher, B. A. J., & Fisher, D. R. (2012). *Techniques of crime scene investigation* (8th ed.). CRC Press.

Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Brooks/Cole.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review, 23*(4), 523–552. https://doi.org/10.1007/s10648-011-9174-7

Gibson, E. J. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585–612. https://doi.org/10.1146/annurev.psych.49.1.585

Kosslyn, S. M., & Koenig, O. (1992). *Wet mind: The new cognitive neuroscience*. Simon and Schuster.

Kundel, H. L., & Nodine, C. F. (1975). Interpreting chest radiographs without visual search. *Radiology, 116*(3), 527–532. https://doi.org/10.1148/116.3.527

McCarley, J. S., Kramer, A. F., Wickens, C. D., Vidoni, E. D., & Boot, W. R. (2004). Visual skills in airport-security screening. *Psychological Science, 15*(5), 302–306. https://doi.org/10.1111/j.0956-7976.2004.00673.x

Mnookin, J. L. (2008). Of black boxes, instruments, and experts: Testing the validity of forensic science. *Episteme, 5*(3), 343–358. https://doi.org/10.3366/e1742360008000440

Nodine, C. F., & Mello-Thoms, C. (2010). The role of expertise in radiologic image interpretation. In E. Samei & E. Krupinski (Eds.), *The handbook of medical image perception and techniques* (pp. 139–156). Cambridge University Press.

Rossmo, D. K. (2009). *Criminal investigative failures*. CRC Press. https://doi.org/10.1201/9781420047523

Russ, A. L., & Safford, B. (2022). Forensic training and cognitive bias: An eye-tracking pilot. *Forensic Science International: Mind and Law, 3*, 100087. https://doi.org/10.1016/j.fsiml.2022.100087

Stainer, M. J., Scott-Brown, K. C., & Tatler, B. W. (2023). Eye-tracking in applied settings: A review of challenges and opportunities for forensic practitioners. *Journal of Applied Research in Memory and Cognition, 12*(2), 155–170. https://doi.org/10.1016/j.jarmac.2023.01.003

Tobii Technology. (2022). *Tobii Eye Tracker 4C user manual*. Tobii AB.

Wilson, M. R., McGrath, J. S., et al. (2011). Perceptual impairment and psychomotor control in virtual laparoscopic surgery. *Surgical Endoscopy, 25*(7), 2268–2274. https://doi.org/10.1007/s00464-010-1546-4

Yarbus, A. L. (1967). *Eye movements and vision*. Plenum Press. https://doi.org/10.1007/978-1-4899-5379-7

Appendix：Supplementary Material

Table S1. Pretest Group Comparison: Key Evidence Zone (Zone_1)

Metric	**Novice M (SD)**	**Prior-Knowledge M (SD)**	***t(49)***	**95% CI**	**Cohen's *d***
TTS (s)	1.873 (1.914)	3.834 (3.943)	-2.362*	(-3.629, -0.2923)	-0.6719
TFD (s)	0.703 (0.866)	1.856 (2.530)	-2.318*	(-2.153, -0.1536)	-0.6596
TFC (count)	2.367 (2.785)	6.238 (8.264)	-2.388*	(-7.129, -0.6136)	-0.6795
PTS (%)	6.399 (5.763)	10.687 (6.378)	-2.503*	(-7.731, -0.8451)	-0.7121
PFD (%)	7.773 (8.786)	9.352 (7.156)	-0.680	(-6.245, 3.0864)	-0.194
PFC (%)	7.450 (7.804)	9.588 (7.388)	-0.984	(-6.505, 2.2280)	-0.280
FFA (s)	6.192 (8.401)	4.226 (4.695)	0.970	(-2.107, 6.0404)	0.2760
FPT (s)	0.391 (0.345)	0.551 (0.488)	-1.528	(-0.371, 0.0504)	-0.435
SPT (s)	1.482 (1.846)	3.283 (3.889)	-2.211*	(-3.437, -0.1642)	-0.6291
RFD (s)	0.384 (0.741)	1.487 (2.480)	-2.302*	(-2.065, -0.1401)	-0.6550
AFD (s)	0.199 (0.154)	0.259 (0.100)	-1.555	(-0.137, 0.0174)	-0.442

Note. Values are presented as means with standard deviations in parentheses. TTS = Total Time Spent; TFD = Total Fixation Duration; TFC = Total Fixation Count; PTS = Percent Time Spent; PFD = Proportion of Fixation Duration; PFC = Percentage of Fixation Count; FFA = Duration before First Fixation Arrival; FPT = First Pass Time; SPT = Second Pass Time; RFD = Revisit Fixation Duration; AFD = Average Fixation Duration; CI = confidence interval. *$p < .05$; **$p < .01$; ***$p < .001$.

Table S2. Pretest-Posttest Gaze Changes in the Prior-Knowledge Group: Key Evidence Zone (Zone_1)

Metric	**Pretest M (SD)**	**Posttest M (SD)**	***t(20)***	**95% CI**	**Cohen's *d***
TTS (s)	2.924 (2.138)	1.375 (0.852)	2.762*	(0.346, 2.753)	0.713
TFD (s)	1.278 (1.255)	0.382 (0.444)	2.44*	(0.109, 1.683)	0.63
TFC (count)	4.267 (3.955)	1.4 (1.639)	2.345*	(0.245, 5.488)	0.606
PTS (%)	11.985 (9.352)	9.443 (5.921)	0.948	(8.29, 3.206)	0.245
PFD (%)	12.792 (18.118)	8.47 (6.105)	0.87	(14.975, 6.332)	0.225
PFC (%)	13.78 (20.263)	8.502 (6.304)	0.952	(17.168, 6.611)	0.246
FFA (s)	4.692 (5.202)	1.025 (2.282)	2.314*	(0.268, 7.066)	0.598
FPT (s)	0.551 (0.43)	0.294 (0.199)	2.186*	(0.005, 0.511)	0.564
SPT (s)	2.373 (2.132)	1.081 (0.746)	2.379*	(0.127, 2.456)	0.614
RFD (s)	0.867 (1.224)	0.154 (0.274)	2.49*	(22.198, 1.653)	0.643
AFD (s)	0.27 (0.095)	0.149 (0.152)	2.981**	(0.034, 0.208)	0.77

Note. Values are presented as means with standard deviations in parentheses. TTS = Total Time Spent; TFD = Total Fixation Duration; TFC = Total Fixation Count; PTS = Percent Time Spent; PFD = Proportion of Fixation Duration; PFC = Percentage of Fixation Count; FFA = Duration before First Fixation Arrival; FPT = First Pass Time; SPT = Second Pass Time; RFD = Revisit Fixation Duration; AFD = Average Fixation Duration; CI = confidence interval. *$p < .05$; **$p < .01$; ***$p < .001$.

Table S3.Pretest-Posttest Gaze Changes in the Novice Group: Key Evidence Zone (Zone_1)

Metric	**Pretest M (SD)**	**Posttest M (SD)**	***t(29)***	**95% CI**	**Cohen's *d***
TTS (s)	5.110 (3.822)	0.743 (1.076)	5.27***	(2.6487, 6.084)	1.099
TFD (s)	1.429 (2.026)	0.265 (0.433)	2.66*	(0.2550, 2.073)	0.554
TFC (count)	4.391 (5.185)	0.957 (1.492)	3.01**	(1.0669, 5.803)	0.627
PTS (%)	28.058 (11.476)	8.612 (9.675)	5.89***	(12.6004, 26.293)	1.228
PFD (%)	23.027 (22.821)	8.307 (13.080)	2.67*	(3.2915, 26.148)	0.557
PFC (%)	22.945 (21.376)	9.119 (13.871)	2.55*	(2.5990, 25.053)	0.533
FFA (s)	2.627 (3.288)	1.043 (2.227)	1.83	(-0.2151, 3.383)	0.381
FPT (s)	0.347 (0.371)	0.133 (0.158)	2.74*	(0.0518, 0.376)	0.571
SPT (s)	4.763 (3.835)	0.610 (1.072)	5.05***	(2.4474, 5.858)	1.053
RFD (s)	1.011 (1.911)	0.114 (0.262)	2.22	(0.0598, 1.734)	0.463
AFD (s)	0.240 (0.125)	0.115 (0.138)	3.16	(0.0428, 0.207)	0.658

Note. Values are presented as means with standard deviations in parentheses. TTS = Total Time Spent; TFD = Total Fixation Duration; TFC = Total Fixation Count; PTS = Percent Time Spent; PFD = Proportion of Fixation Duration; PFC = Percentage of Fixation Count; FFA = Duration before First Fixation Arrival; FPT = First Pass Time; SPT = Second Pass Time; RFD = Revisit Fixation Duration; AFD = Average Fixation Duration; CI = confidence interval. *$p < .05$; **$p < .01$; ***$p < .001$.