Statistics for Forensic Analysts

a study validating the use of on-line micro learning to increase the statistical literacy of digital / multimedia forensic analysts

Jim Hoerricks, PhD
Presented to the Teachers College Faculty
of Western Governors University, Salt Lake City, UT, June 2019

Abstract

In a small-scale mixed methods action research study, the problem of the statistical literacy of analysts engaged in the digital / multimedia forensic sciences was investigated. Specifically, the study sought to answer how an online instructional unit, which includes forensic analysis techniques, delivered as micro learning affects achievement of participants as measured by a pre/post assessment? The results found that there was a 29% increase in the median score for the group, moving from a pre-test median score of 59% to a post-test median score of 88%, indicating that the intervention was successful. Not only did all of the learners pass the 75% threshold needed to pass the post-test, but all improved their scores rather dramatically after having experienced the intervention.

Topic

A significant gap exists in the administration of justice in the United States in that a majority of digital / multimedia forensic science practitioners (forensic analysts) lack a basic knowledge of statistics as well as general inferential and probabilistic reasoning for forensic science applications (Houck, 2013). Forensic science is the systematic and coherent study of traces to address questions of authentication, identification, classification, reconstruction, and evaluation for a legal context, with a trace being any modification, subsequently observable, resulting from an event (OSAC, 2017). In adjudicated cases, the prosecution seeks to link trace evidence to a specific individual or group of individuals alleged to have committed a specified violation of the law. With this in mind, inaccuracy in interpreting results has resulted in the conviction of many who have later been found to be factually innocent (Gabrielson, 2009a, 2009b).

Given the identified learning deficit (Saks, 2005), the costs involved in training practitioners to competency (UTA, 2018), and the available infrastructure to facilitate such a training / education scheme, a new model of statistical education for forensic science practitioners – on-line micro learning - is proposed. The objective is to equip forensic analysts to become responsible producers and discerning consumers of statistics and confident exponents of elementary probabilistic reasoning (Aitken, et al, 2010). But there are problems to overcome.

Micro learning, an electronic style distance learning (e-learning) approach has experienced year over year growth since the delivery of the first on-line course. This increased enrollment in online courses can be traced to a variety of factors. These include a greater choice of programs and coursework (as opposed to being limited to what's available locally); higher levels of comfort with the technology (the proliferation of high-speed internet has made the user experience more satisfying); as well as the convenience and flexibility that on-line programs offer (Hill, 2007). Yet, in spite of this, the traditional model of in-person training that is developed and hosted within the organization remains the preferred option for police agencies in the United States. Each agency has their own training office, often coordinating with a regional or state-level standards-setting body for curriculum development assistance. These standards-setting bodies are often staffed by former local law enforcement trainers / supervisors who aid in the implementation of training initiatives state-wide (CA POST, 2019).

Whist this arrangement may work for traditional policing topics, it's woefully inadequate for training and education in the sciences. As police agencies move beyond patrol functions and basic detective work into the forensic sciences, the process of taking a police officers and turning them into forensic scientists is arduous and expensive when handled internally. Yet, most agencies work in this way – attempting to learn by doing – and not engaging with subject matter experts in the sciences. A recent series by ProPublica notes how this formula of in-house training and validation can go wrong, perverting the cause of justice in the process (Gabrielson, 2019a, 2019b).

Additionally, this insular system does not lend itself to innovation. Training and education in law enforcement often lags behind it's corporate peers in terms of flexibility in spending and program options. Furthermore, spending on training is often skewed towards the higher ranks. Hall (2000) noted that spending on continuing education and training for senior management exceeded $1000.00 per learner in 68% of the organizations surveyed whilst the majority of spending on the lowest ranks was at the $100.00 or less level. Given that specialist services, like forensic science, require a great deal of training in order to begin working, as well as continuing education options that keep pace with changes in the law and technology, this spending pattern is recipe for failure.

Currently, instructional delivery in the forensic sciences happens in one of two locations. Primarily, agencies prefer to bring instructors and courses to their location. Training sessions can happen in roll-call / briefing rooms, conference rooms, or training centers. This tends to be the most cost-effective option when agencies need to train large groups of employees. However, there are not enough instructors to meet the needs of the over 18,000 US police agencies. Additionally, given the learners' proximity to the job site, circumstances often conspire against their full attention to the learning experience. Occasionally, the learner will be called out of the training session to perform their normal work duties. Secondarily, agencies will send employees to off-site training courses. This option is popular when an agency has less than five employees to train. There is an advantage to the learner in going off-site as distractions can be minimized by the training provider. Yet, as travel costs rise, this type of spending is facing increasing scrutiny (OMB, 2013).

Optimally, the learning context should be divorced from location. Micro learning, delivered on-line, is a learner-centric approach that offers just-in-time training that is accessible by the learner where and when the learner is available (Pandey, 2018). Micro learning is affordable and agile, has a shorter development cycle, is easier to update, has a wider potential application, and a higher impact on learning and application versus traditional on-premises training delivery (Pandey, 2018).

Bombarding the brain with an abundance of data, as often happens in a classroom environment or by studying a large document ahead of a high-stakes exam, is not an optimal way of delivering technical training to a mature adult learner population. Micro learning provides small segments of information to the learner so that their mind can process it more effectively, assisting in better retention. Thus, adult learners are able to absorb the information before they have to move onto the next activity or module. The learner can also decide when and where they participate in the training experience, which allows them to pick a time and place that are conducive to the training process. For example, instead of having to sit in a crowded conference room that has a variety of distractions, they can wait until they are in a quiet place where they can simply focus on the training activity (Pappas, 2018). This becomes particularly important given the amount of knowledge, skills, and abilities that a person needs in order to begin their work within the forensic sciences (SWGDE, 2010).

Problem Statement

The research problem that was investigated is the suitability of utilizing micro learning to address the lack of knowledge about statistics within the digital / multimedia forensic analyst community. In the current state of affairs, a majority of digital / multimedia forensic analysts and a minority of forensic analysts in the physical sciences are not proficient in this fundamental aspect of their work (NRC, 2009). A review of the available literature illustrates this gap, as well as how this deficiency has led to the false convictions of countless individuals. Instruction is thus required to bring the population of analysts to competency.

Problem Background and Causes

The National Research Council outlined (NRC, 2009) the need for education and training by forensic scientists. Chapman (2015) notes that many education and training requirements in the judicial system have no specificity. If a practitioner is required to earn 40 credit hours of continuing education within a given 5-year period, there is no specification as to what those hours should cover. Additionally, there is very little enforcement as to actually receiving continuing education. The US state of Texas is the only place where one will find a specific educational requirement for statistics amongst forensic science practitioners (Hervey, 2017). The Texas initiative comes from it's new licensure scheme for practitioners where statistics in a general sense is one of 6 tested domains.

As an area for study, Karie & Venter (2015) note that instruction on basic empirical research is needed. Analysts lack basic understanding on conducting experiments, controlling variables, selecting appropriate samples, calculating the appropriate sample sizes, and even working within a basic plan. Along these lines, Neufeld (2005) noted that most experiments were conducted for actual case work, and as such, very little validation or validation studies exist for the work that is performed in the forensic sciences. If Saks, et al, (2016) and Vincze (2016) are to be believed about the state of forensic practice, a solid curriculum design starts to take shape that includes not only the foundational study of descriptive and inferential statistics, but also the fundamentals of probability and inferential reasoning. Whilst just about every college on earth has a basic statistics course on it's schedule, there is as yet no course offered that focuses on the unique needs and context of forensic science.

Analysts working in the natural sciences often face educational prerequisites as a condition of employment. Most will have taken some form of quantitative analysis or statistics in college. However, analysts working in the physical sciences, such as tool-mark analysts, latent print analysts, or digital / multimedia analysts, often arrive in these positions without a college degree. In state / local police service, it is often credentialed peace officers who are filling these roles. As such, they may not have an exposure to education in statistics or probability. Yet, these sciences form the foundation of many of the conclusions reached in the physical sciences. If the role of the forensic scientist in the court room, as expert witness, is to inform the Trier of Fact, and the so-called expert has no grasp of the underlying science or even that their work is flawed, then problems can cascade (Eastwood, Caldwell, 2015).

In terms of entry-skills, or skills that analysts must have already mastered that are associated with a learning goal, courses in statistics tend to feature higher level maths as a prerequisite. An example of this can be seen in the current course schedule for the University of Texas at Austin (UTA, 2018). In order to register for Introduction to Statistics and Probability, undergraduate students must have successfully completed a three-course series covering differential and integral calculus. The curriculum and instructional designers in the UT system believe that a firm grasp of calculus is a necessary entry-skill for their introductory level statistics and probability class. Combine the high bar of entry skills with the fact that the learner must apply for and be accepted to the university in order to sit for the course, and the level of complexity in solving the problem escalates. As regards prior knowledge, the problem to overcome are the many misconceptions that surround the topic of statistics. Without guidance, analysts will be left to interpret the new content in light of the associations they can make with their misconceptions on the topic (Dick, et al, 2014). The new information will be layered upon a foundation comprised of assumptions and partial knowledge. Thus, the goal of a program will not be met.

With these problems in mind, the Trier of Fact is now expecting analysts to express their findings in terms of likelihood ratios and confidence intervals (Martire, et al, 2014). Many analysts do not grasp the fundamental issues within the calculation of such values, but offer their best guesses (Spellman, 2018) in their reports and during their testimony (Thompson, 2018). Is the analyst testifying as to an expression of the subjectivity and uncertainty associated with forensic science evidence, or the strength of the analyst's opinion (Martire, et al, 2014)? Few can adequately explain the nature of likelihood ratios. Given the over 90% plea rate in the justice system, few have reason to (Glazebrook, 2018).

Research Methods

This study featured a sequential Quan → Intervention → Quan mixed methods action research design, which consisted of three chronological strands. The study began with an assessment of the learners' initial knowledge of statistics, followed by the delivery of the course (the intervention), and concluded with a summative assessment to test the learners' knowledge of statistics after the intervention had been delivered.

In the initial assessment, the quantitative data were collected and analyzed to inform the intervention. Then, quantitative data were collected to test the success of the intervention (Ivankova, 2015) via the summative assessment. The primary purpose of this design was to use data to drive action and improve outcomes in the forensic sciences.

The learner group (N=7) were given a pretest within the LMS. The test questions, listed in Appendix B, were divided into sections in order to correspond with the subjects within the intervention. The sections were, Empirical Research (3 randomized questions from a pool of 6 questions), Descriptive Statistics (3 randomized questions from a pool of 6 questions), Correlational Statistics (3 randomized questions from a pool of 7 questions), Inferential Statistics (3 randomized questions from a pool of 14 questions), Probability (1 question from a pool of 1 question), and Statistics for Forensic Science (4 randomized questions from a pool of 10 questions). The presentation of the 17 total questions on the pre-test were randomized such that no two learners were presented with the same questions, in the same order.

The learners were then provided with access to the intervention, a course of instruction covering statistics for forensic analysts. The syllabus, linked here, illustrates the path through the coursework. The coursework, though delivered on-line as micro learning, is equivalent to an eight-hour course of instruction. Thought the learners were given 10 days to complete the intervention, all were finished in less than 7 days.

The post-test was administered from within the LMS immediately following the course of instruction. The structure of the test was identical to the pre-test, both sharing the same question pool and groupings of questions. Given the randomizing of the questions, no learner faced the exact question assortment or arrangement that was faced in the pre-test.

The results indicate that the intervention was successful in addressing the research question, how does an online instructional unit which, includes forensic analysis techniques, affect achievement of participants as measured by a pre/post assessment? The results, explored in detail below, showed that no learner passed the 75% threshold of correctly answered questions on the pre-test. After the administration of the intervention, all learners passed the 75% threshold of correctly answered questions. Overall, there was a 29% increase in the median score for the group, moving from a pre-test median score of 59% to a post-test median score of 88%.

Conclusion

A small group of analysts assisted in this enquiry; investigating the fitness of an online unit of instruction in improving their statistical literacy. The results of the pre-tests reinforced what was discovered in the literature review, that the gap in knowledge exists. The results of the post-test, by a diverse group of learners, illustrated that on-line micro learning can be effective in bringing analysts to competency in statistical literacy. That the improvement was rather uniform speaks to the appropriateness of the level of complexity of the course of instruction.

The results found that there was a 29% increase in the median score for the group, moving from a pre-test median score of 59% to a post-test median score of 88%, indicating that the intervention was successful. Not only did all of the learners pass the 75% threshold needed to pass the post-test, but all improved their scores rather dramatically after having experienced the intervention.

Given the number of analysts currently employed in the whole justice system (employed in public and private service) and the amount of cases before the courts in the US that require their services, the problem of a statistically illiterate workforce seems monumental. The legacy training paradigm of in-class sessions with live instruction is no longer fit for purpose in addressing this issue. With this massive problem in mind, the study's results indicate that the problem of a lack of statistical literacy could be mitigated by an on-line course of instruction, delivered as micro learning. Addressing the problem in this fashion can help improve the current situation whilst keeping the costs down.

With the validation of the concept complete, the full course of instruction has been released to the general public. It can be found on the researcher's website (link).

About the researcher

Jim Hoerricks, PhD, is a Certified Audio/Video Forensic Analyst (AVFA), is the author of the best-selling book Forensic Photoshop (available on Amazon.com), is a co-author of Best Practices for the Retrieval of Video Evidence from Digital CCTV Systems (DCCTV Guide), published by the Combating Terrorism Technical Support Office (CTTSO), is retired from police service where his unit was responsible for the collection and analysis of digital / multimedia evidence (LAPD's Electronics Unit), and currently serves the Organization of Scientific Area Committees on Forensic Science (OSAC) as the Video / Image Technology and Analysis (VITAL), Video Task Group Chair.

References

Aitken, C., Roberts, P., & Jackson, G. (2010). Fundamentals of Probability and Statistical Evidence in Criminal Proceedings [PDF]. London, UK: Royal Statistical Society.

Chapman, L. (2015). Playing forensic science monopoly. Australian Journal of Forensic Sciences, 47(4), 402. Retrieved from https://wgu.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=109574957&site=eds-live&scope=site

Dick, W., Carey, L., Carey, J. (2014). Systematic Design of Instruction, The, Loose-Leaf Version, 8th Edition.

Eastwood, J., & Caldwell, J. (2015). Educating Jurors about Forensic Evidence: Using an Expert Witness and Judicial Instructions to Mitigate the Impact of Invalid Forensic Science Testimony. Journal Of Forensic Sciences, 60(6), 1523–1528. https://doi-org.wgu.idm.oclc.org/10.1111/1556-4029.12832

Gabrielson, R. (2019a). The FBI Says Its Photo Analysis Is Scientific Evidence. Scientists Disagree. Retrieved from https://www.propublica.org/article/with-photo-analysis-fbi-lab-continues-shaky-forensic-science-practices

Gabrielson, R. (2019b). FBI Scientist's Statements Linked Defendants to Crimes, Even When His Lab Results Didn't. Retrieved from https://www.propublica.org/article/fbi-scientists-statements-linked-defendants-to-crimes-even-when-his-lab-results-didnt

Glazebrook, J. S. (2018). Miscarriage by Expert. Victoria University of Wellington Law Review, 49(2), 245–258. Retrieved from https://wgu.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=lft&AN=131890309&site=eds-live&scope=site

Hall, B., LeCavalier J. (2000). E-Learning Across the Enterprise: TheBenchmarking Study of Best Practices.[Online] Retrieved from http://go.brandonhall.com/l/8262/2019-02-18/965dzp

Hervey, B. P. (2017). A Texas Approach to: The Intersection of Science and the Law. Judges' Journal, 56(4), 20.

Houck, M.M. (2013). Intellectual infrastructure: a modest critique of forensic science. Science & justice: journal of the Forensic Science Society, 53 1, 1. Retrieved from https://ac-els-cdn-com.ezproxy.trident.edu/S1355030613000026/1-s2.0-S1355030613000026-main.pdf?_tid=51f1ff91-47b3-4983-bc37-47163e5b659e&acdnat=1546984525_4093f861a0e51771d30fb7dc535100f5

Ivankova, N. (2015). Mixed methods applications in action research. SAGE Publications.

Karie, N. M., & Venter, H. S. (2015). Taxonomy of Challenges for Digital Forensics. Journal of Forensic Sciences, 60(4), 885. Retrieved from https://wgu.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=108611354&site=eds-live&scope=site

Martire, K. A., Kemp, R. I., Sayle, M., & Newell, B. R. (2014). On the interpretation of likelihood ratios in forensic science evidence: Presentation formats and the weak evidence effect. Forensic Science International, 240, 61–68. https://doi-org.wgu.idm.oclc.org/10.1016/j.forsciint.2014.04.005

National Research Council, Committee on Identifying the Needs of the Forensic Sciences Community. (2009). Strengthening Forensic Science in the United States: A Path Forward. Retrieved from https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf

Neufeld, P. J. (2005). The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform. American Journal of Public Health, 95, 107–113. https://doi-org.wgu.idm.oclc.org/10.2105/AJPH.2004.056333

Office of Management and Budget. (2013). OMB Instructs Agencies To Trim Travel And Conference Costs. Retrieved from https://www.fedmanager.com/featured/9-general-news/739-omb-instructs-agencies-to-trim-travel-and-conference-costs

Organization of Scientific Area Committees on Forensic Science. (2017) A Framework to Harmonize Forensic Science Practices and Digital/Multimedia Evidence. Organization of Scientific Area Committees on Forensic Science Task Group on Digital/Multimedia Science.

Pandey, A. (2018). 10 Benefits of Microlearning-Based Training. Retrieved from https://elearningindustry.com/10-benefits-microlearning-based-training

Pappas, C. (2018). Microlearning In On Training: 5 Advantages And 3 Disadvantages. Retrieved from https://elearningindustry.com/microlearning-in-online-training-5-advantages-and-3-disadvantages

Saks, M. J., Albright, T., Bohan, T. L., Bierer, B. E., Bowers, C. M., Bush, M. A., … Imwinkelried, E. (2016). Forensic bitemark identification: weak foundations, exaggerated claims. Journal of Law & the Biosciences, 3(3), 538. Retrieved from https://wgu.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=120820446&site=eds-live&scope=site

Saks, M., & Koehler, J.J. (2005). The coming paradigm shift in forensic identification science. Science, 309 5736, 892-5. Retrieved from https://pdfs.semanticscholar.org/820b/41b12107ea34a54e32243b9e1f1980ca256e.pdf?_ga=2.219044569.119495570.1546982617-1361324278.1546982617

Scientific Working Group on Digital Evidence. (2010). SWGDE/SWGIT Guidelines & Recommendations for Training in Digital & Multimedia Evidence. Version: 2.0 (January 15, 2010). Retrieved from https://swgde.org/documents/Current%20Documents/SWGDE-SWGIT%20Guidelines%20and%20Recommendations%20for%20Training

Spellman, B. A. . (2018). Communicating Forensic Evidence: Lessons from Psychological Science. Seton Hall Law Review, 48(3), 827–840. Retrieved from https://wgu.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=lft&AN=130345067&site=eds-live&scope=site

UTA. (2018.). Courses: Undergraduate Course Inventory. University of Texas, Austin. Retrieved from https://stat.utexas.edu/undergraduate/courses-undergraduate#introprobnstats

Thompson, W. C. . (2018). How Should Forensic Scientists Present Source Conclusions? Seton Hall Law Review, 48(3), 773–813. Retrieved from https://wgu.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=lft&AN=130345071&site=eds-live&scope=site

Vincze, E. A. (2016). Challenges in digital forensics. Police Practice & Research, 17(2), 183. Retrieved from https://wgu.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edb&AN=112454861&site=eds-live&scope=site