JOURNAL OF INTEGRATED OMICS

A METHODOLOGICAL JOURNAL

HTTP://WWW.JIOMICS.COM

REVIEW | DOI: 10.5584/jiomics.v8i1.215

Quality Improvement for Criminal Investigations: Lessons from Science?

Hans Ditrich *

Institute for Science and Research - .SIAK - Dept. I/9. Federal Ministry of the Interior.

Received: 13 November 2017 **Accepted:** 4 March 2018 **Available Online**: 29 March 2018

ABSTRACT

Criminal investigations generally aim at discovering previously unknown facts. The same is true for scientific (or academic) research. Both follow a rather tight framework of rules – most importantly, the principles of objectivity, reliability and validity. However, some of the intentions differ. Science generally attempts to discover and/or explain new principles, while criminal inquiries are instead usually bound to past, often singular, events. For example, the methods used in forensic investigations are required to be well established, standardised and undisputed inasmuch as possible. In contrast, the exploration of new methods is an important feature of the advancement of science. Consequently, both tendencies – similarities and opposites – can be discerned when comparing criminal and academic examinations.

The 'Pareto principle' indicates that the vast majority of all criminal investigations run rather unproblematically. Nevertheless, the highest quality criteria must be guaranteed for these and the remaining, more challenging cases as well – based on the 'fair trial' principle. Acknowledging that mistakes are inevitable (Murphy's law), methodical approaches for error identification, handling, management and reduction are essential.

Error correction mechanisms that are typical for forensic statements normally include a second source of expertise and/or an appeals procedure. In academic science, however, the peer review system has long been established as the most important quality control and error correction system. Furthermore, possible mistakes can usually be corrected in later, more detailed studies. However, the central position of forensic experts and criminal investigators in a legal procedure and the severe personal consequences of incorrect statements emphasize the high importance of continuous improvement of both the qualifications of the investigators and the quality of their methods.

Nevertheless, error reduction provisions should not be restricted to technical measures such as quality management and accreditations. Furthermore, a systemic/organisational approach towards error management seems promising. This involves, among other measures, a systematic examination of mistakes and the recognition of the human factors that underlie them. Nevertheless, an indispensable component for quality enhancement is intense cooperation from both sides – the criminalistic and forensic practice as well as the scientific (basic) research .

Keywords: Forensic, Science, Quality, Expert, Testimony, Error

1. Introduction

According to the general humanitarian principles and the European convention on human rights [1], every suspect is entitled to a fair trial and an investigation that is based on objectivity. The latter, however, is a target that is often difficult to meet, since human decisions are often biased by motives that are not evidence-based. Thus, in a tribunal scenario, objectivity is often anticipated from scientific investigations, usually in the form of forensic science reports and/or testimony from expert witnesses. While this input certainly increases the objectivity, reliability and validity of decisions, the significance of scientific results is frequently misjudged (i.e., overestimated) by laypersons. Therefore, it may be useful for judges, jury members, criminal investigators as well as for the experts, to recall a few principles from the theories of science to gain a better perspective for scientific statements.

***Corresponding author:** Hans Ditrich, Institute for Science and Research - .SIAK - Dept. I/9. Federal Ministry of the Interior Herrengasse 7; 1014 Vienna, Austria. Tel: +43-1-53126-4828; E-Mail Adress: hans.ditrich@bmi.gv.at .

2. Scientific Framework

From the numerous philosophical concepts that interpret the process of human cognition, critical rationalism [2] seems most appropriate as an initial point for scientific investigations. It accepts the world as a factual entity (i.e., things exist independently from the observer) and acknowledges the limitations of the human mind (and instruments) to fully perceive all features of reality. Furthermore, this system proposes a method to gain knowledge from errors, the latter being accepted as an inevitable component of all heuristic explanation efforts.

Thus, the scientific process generally involves the formulation of hypotheses that are checked with experiments that are designed to find errors in the proposed model. Eventually, the theory is refined to circumvent the contradictory findings and an improved postulate is then further tested (Figure 1). Note that the explanation is modified and not the experimental results – the latter are regarded as valid unless proven otherwise.

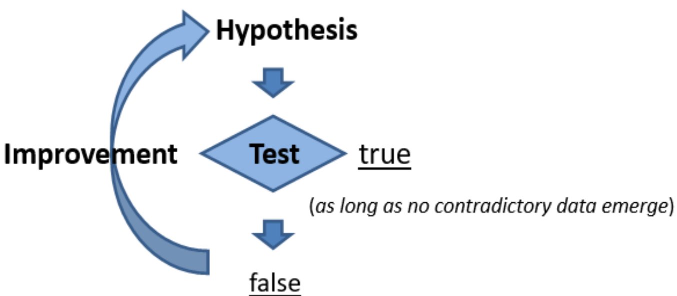

Figure 1. Schematic flow of decision making according to the critical realism concept.

Successively, the refined theories from this process can eventually be regarded as "true" and may be absorbed in the corpus of formal scientific knowledge – unless they are falsified by a later experiment.

It is obvious that this procedure is not appropriate for forensic/criminalistic purposes. It would be rather questionable to proceed until future studies confirm or dismiss a finding that is the basis for a verdict. However, this may be the best chance for revealing misjudgments and exoneration [3] of the accused.

There are two possible types of reasoning in the interpretation of results. The more convenient way for law enforcement (and in many other situations) is reductionist (linear or deterministic) reasoning. Thus, the answer of a questioned statement can only have two possible outcomes – yes or no, true or wrong, guilty or innocent, etc. An example would be a testimony such as: 'Person X was not at the crime scene as he or she was in custody at the specified time'. This applies (only) to one single, specific event. Unfortunately, questions that can be answered with a reductionist approach are rather rare.

Frequently, the answer to a question or the outcome of an investigation, measurement, etc., is a stochastic value – the product of a probability distribution. Such statements normally meet the criterion of a defined standard error (see below). However, in order to minimize that error, a large (or even infinite) number of coherent observations would be ideal. The latter is obviously problematic with forensic samples that are often small and/or cannot be collected in a standardized way. An example for this type of result would be: 'The DNA sample from the crime scene is from person X with an error rate of $p < 0.001\%$'.

By convention, for describing the results, the zero-hypothesis is applied, e.g., causality does not exist, an event did not happen, two actions are not related, a person is not guilty, etc. The value 'p' gives the probability of this assumption being wrong, where $p = 0$ means that a result is certain, $p = 1$ indicates that the statement is false. The values $p < 5\%$ (often also given as $p < 0.05$) is 'significant', $p < 1\%$ (or $p < 0.01$) is 'very significant' and $p < 0.1\%$ ($p < 0.001$) would be 'highly significant'. Accounts of this type are highly appreciated by juries; however, several caveats apply (and should be explained by the expert): Statistical data are prone to a number of methodological errors, e.g. clustering effects, sampling bias, etc. These effects can be avoided by taking a sufficient number of homogenous samples. However, forensic traces are seldom homogenous and usually limited in number. Furthermore, an answer based on statistical results is valid for its statistical basis and not necessarily for each single item (outlier). Also note, that a probability of 1 : 10 000 000 ($p= 0.0000001$) – e.g., in a DNA assessment – still implies that there are at least fifty persons in the EU (out of approx. 511 M citizens) with that combination of features. Thus, relative numbers can provide a wrong impression (compare also [4]).

Furthermore, statistics can only prove correlations and not causality. Thus '*cum/post hoc ergo propter hoc*' (together with 'X'/following 'X', therefore 'Y') errors are likely to appear. The display of statistical data may also lead to an incorrect impression of 'scientific certainty', since all cognitive and methodical errors that might have happened before are masked by the perceived assertiveness of the result.

In addition to the desirable outcomes of a measurement – i.e., an assumption being confirmed or rejected – there is also the possibility that a result may be erroneous (Fig. 2). This can happen for a large number of reasons – e.g. methodical deficiencies, human flaws, or even intentional deception (fabricated evidence).

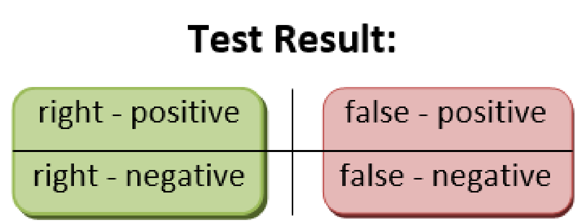

Figure 2. Diagram of the theoretically expectable results of an analysis.

Thus, beside a correct result – a searched feature is present or not – the possibility that a test may be inadequate and thus unable to detect the questioned feature even if it is present must be considered. Vice versa, a positive result may also be derived from a number of sources of error such as oversensitive test systems, contamination, high background "noise", etc. The human factor obviously influences both types of false outcomes. Reduction of these errors is the subject of quality assurance programs, but a complete prevention of all mistakes is theoretically impossible and practically unfeasible.

It should be noted, however, that the rather tight set of rules previously outlined is primarily followed in natural sciences and also in medical research. Humanities, such as psychology, sociology, criminology, etc., are often less strict in their interpretation of epistemological principles

3. Methodological Aspects

Minimum requirements were established attempting to insure a certain degree of relevance and reliability of investigation results in the USA. The Frye standard [5] stated that a method for investigation should be 'accepted by practitioners' – a concept that is problematic for several reasons (it should not be forgotten, that e.g. medical 'practitioners' insisted that e.g., bloodletting was a useful therapy for a long time). Consequently, the Frye standard was replaced by the Daubert standards [6] in 1993. The latter expect that, in addition to being accepted by the 'scientific community', a valid investigation method should be testable, published in a peer reviewed journal and give due regard to the standard error of the results (see also [7]).

Scientific and criminalistic investigations are not equal in their general objectives. While science basically aims to identify general principles, specific questions about single events, i.e. individual crimes, are usually targeted by the various forensic disciplines. For example, the scientific principle of reliability (and reproducibility) demands that a result should be repeatable, obtained under controlled conditions and statistically secured. Consequently, all parameters should be under the control of the experimenter (at least theoretically), a condition that is impossible to achieve in criminal investigations. In forensic practice, traces are scarce, often poorly preserved, sometimes contaminated and therefore not 'standard' in many respects. However, some academic disciplines like palaeontology or pathology may also have similar problems – being restricted to material that is limited or even unique.

Furthermore, while it is generally the aim of scientific research to find new, previously unknown principles – forensic findings should not primarily be innovative, but be widely acknowledged instead. This is a main difference between academic science and forensic investigations. Additionally, science aims usually at discovering and explaining general principles, while forensic work is essentially connected with the clarification of past, normally singular events (i.e., crimes).

The potential consequences of expert testimony imply the demand for the highest professional standards for the quality of an investigation as well as for the competence of the examiners. However, errors inevitably occur in spite of all efforts to avoid them.

The traditional error correction mechanism in academic science is the peer review process (and thus is requested by the Daubert standards). However, reviewed studies may also be wrong [8] (compare also the discussions following the publications of Sokal [9] or Lindsay and Boyle [10]). A second level of validation in academic science is derived from follow-up studies that should be able to replicate the previous results (at least in theory).

This system established in the scientific process is clearly inappropriate for criminalistic/forensic/juridical statements, since it requires an abundance of time and a generalizing approach to a question that is often not adequate for single cases. Instead, the established system involves appeal and obtaining a second expert opinion. Note however that these steps also require time and often a certain financial capacity of the culprit.

Numerous sources of mistakes exist in addition to the potentially false test-results mentioned above. Quality management systems (e.g. ISO/IEC 17025:2005) attempt to insure adequate technical standards and analytical procedures. These rules propose uniform solutions for essentially identical tasks, thus eliminating variability inasmuch as possible. Nevertheless, it is essential to check first if the questioned issues indeed are identical and if all preconditions for a valid analysis are met. Both are rather difficult to assess for the defendant or lawyer. It should also be noted that an implemented SOP (standard operating procedure) or GLP (good laboratory practice) can only assure quality on a formal basis – by minimizing deviations from an 'ideal' prototype/standard – thus eliminating all variations, experimentation and creative processes. Consequently, these frameworks are primarily useful in routine production or analysis sequences (Fig. 3), but may be inherently obstructive for innovative scientific discovery. Nevertheless, in several forensic fields (e.g., in drug- or DNA -analysis) standardisation and quality management routines are certainly indispensable. Alternatively, investigations in exceptional or complex criminal cases can demand flexible and inventive approaches that are not formally standardized or even require the use of (yet) unpublished methods.

4. The Human Factor

Not surprisingly, a main source of flaws originates from human imperfection. Several factors contribute to wrong conclusions, ranging from inadequate crime scene work to basal human cognitive mechanisms (e.g., [11]). Human perception is neither a camera nor a computer and thus prone to flaws in perceiving and drawing conclusions. Other shortcomings, also often resulting from limited personal,

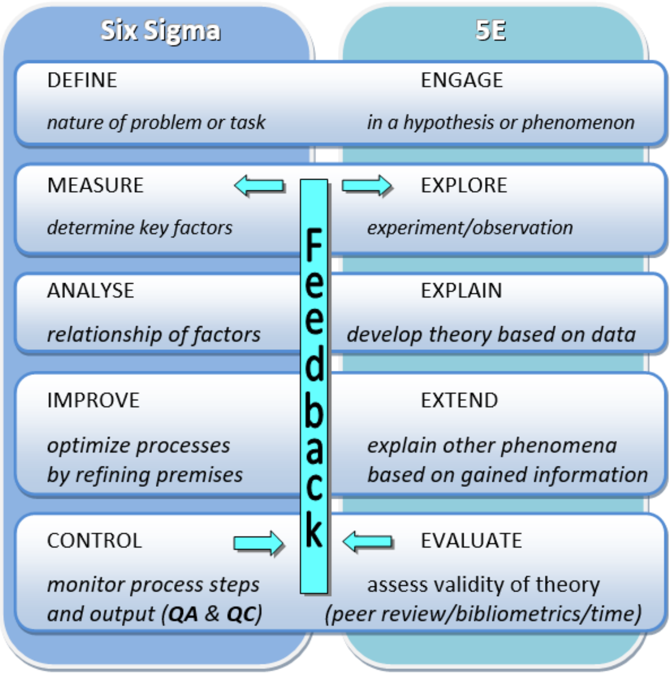

Figure 3. Comparison of quality improvement programs for production/services (Six Sigma) and academy/education (Five E). Note that specific quality assurance (QA) and control (QC) are the last steps in the production schedule while analogous steps in the scientific procedure are less explicit.

equipment or time may adversely influence investigations. However, the latter difficulties may at least partly be countered by organisational and training measures (e.g., [12]).

Encouragingly, the main source of erroneous verdicts is probably not expert statements but wrong witness reports. About 60% of wrong convictions in review of US-courts verdicts [13] result from false testimony (mainly misidentifications – see also [14]). Additionally, mistakes by authorities (police, prosecutor, jury), false accusations (e.g., by a snitch) and mistakes in defence (lawyer) as well as false confessions were found to be major contributors, besides wrong results from forensic investigations [15].

A further problem that is not restricted to formally controlled or accredited processes but rather refers to many types of reporting of results is connected with interpretation. The narrative element, i.e., communication between expert and jury, either in a written report or in testimony at court is an often neglected factor. The inevitable differing conceptions and sometimes even language of practitioners and tribunal must actively be countered by both sides. Formalized schemes of responses [16, 17] such as: '… is (very) likely', '… cannot be approved or denied', '… is (very) unlikely', etc., which are sometimes linked with probability values in reports, may be facilitating conventions but do not support a deeper understanding of a statement (compare e.g., [18, 19]. Instead they attempt to fill a gap between something 'is' or 'is not' (taking into account methodological limits and error margins). If such arbitrary scales are secondarily related to numbers, the results become even more indefinite. For example, is a result of 2.5 in the scale mentioned before (between … is likely and … cannot be approved or denied) really 'better' than a 4 (… is unlikely)? The frequently demanded comparability of testimony from different experts is only seemingly provided by a choice of pre-defined diagnoses, especially when the results strongly depend on the opinion of the expert.

Investigations that are heavily dependent on interpretation by a practitioner like e.g. fingerprints (dermatoglyphics), handwriting, forgery, etc., but also many other kinds of expert assessments (medical, psychiatric, etc.) are principally based on the individual competence of the respective specialist. Formal accreditation procedures or controlling schedules cannot fully assure the correctness of testimonies in such fields (e.g., [20]). Investigation routines based on the four-eyes principle and eventually a second, independent expertise can at least partly overcome this problem (compare also [21]). However, as long as there are expert witness reports used in court they will remain open to criticism. Ensuring, improving and maintaining the qualification of experts and their reports is an essential responsibility also for academia.

Several academic disciplines are less based on observation and rely more on theories. This is typically the case for human and social sciences, psychology, economics, art, etc. However, both systems – data driven and more intuition-based – are commonly applied in these fields. Accordingly, such reports can be applied to cases of law, although with special caution as e.g., the Daubert criteria might not be fulfilled.

Even if the highest professional standards can be maintained, certain errors rooted in the human cognitive system are difficult to control. Cognitive dissonance gives a subconscious bias towards the explanation that causes the least emotional conflict, thus clearing the mind from strong antagonistic sensations. An example for this phenomenon is the 'neutralisation' that criminals exhibit with respect to their victims, (i.e. the victims provoked the attack, have only self to blame, etc.). However, the interpretation of evidence by an investigator underlies similar cognitive mechanisms. This can result even in the inability to perceive adverse details once a theory has formed (i.e. after a few seconds). Precognition of seemingly or factually unrelated details can strongly influence the collection and subsequent evaluation of evidence (e.g., [22, 23, 24, 11]) and external influences (including the judge – compare [25]) can severely affect the outcome of a lawsuit.

Cognitive bias affects everyone in daily life. However, in criminal investigations it may lead to serious adverse effects. Experienced criminal investigators considered in a survey selective perception/expectation/confirmation bias, anchoring/'pars pro toto' errors and 'onus probandi' infringements (shifting the burden of proof to the suspect) as those cognitive factors most likely to negatively affect criminal investigations [12].

Acknowledging the fact that 100% error free conditions are very difficult to attain and impossible to maintain, poses the problem of error handling strategies. As mentioned already, in individual cases an appeal and a second expert opinion can correct a wrong expert assessment. Quality management systems can generally reduce the error rate in routine processes. Nevertheless, errors will occur.

Mistakes that happen in spite of these efforts should strictly be investigated (instead of the understandable tendency to cover them up). In several fields where errors can have dramatic consequences, like e.g., in air transportation or also in clinical medicine, painstaking investigations are often carried out after a disastrous event. The aim of these studies is not only to investigate the specific incident, but also to refine the rules and procedures in order to prevent future adversities. Institutional boards of inquiry with external experts could identify 'hot spots' in forensic investigations by systematic analysis of errors with the aim of suggesting ways to avoid or improve such pitfalls [12].

The demand for objectivity in juridical procedures is often symbolized by the blindfolding of the Roman godess 'Justitia'. Organizational measures, e.g., anonymizing samples, double blind tests, four eyes principle, etc. are important steps for supporting this principle but most important are the appropriate attitude and ethics of the investigator (among many others, see also [26]). Unfortunately, these assets are difficult to teach and to surveil. The integrity of officers is strongly dependent on both the individual values and also on the organizational culture.

5. Conclusions

As indicated before, the importance of forensic expert testimony in the legal system implies the need for highest standards for both the qualification of an expert and the level of technical/methodological quality. The best available practices must be applied to every criminal case – a request that is based on the 'fair trial' principle. The general 'scientific' rules of objectivity, reliability and validity form a common basic framework for academic science as well as forensic investigations (while also maintaining ethical and economical standards). Requirements to fulfil these demands are the continuous professional training of the experts, application and further improvement of technical standards (best practice, GLP, SOP, etc.) and the development and implementation of a system for error management.

While extensive formal education is an essential prerequisite for forensic work, it is evidently the everyday challenge of practical operation that provides the demand as well as the criteria for applicability and validity for all forensic methods. It is common knowledge that practical application is the ultimate test bench for academic results. While it is probably rare that the same person is active equally in scientific research as well as in practical criminalistic work it is essential to provide an environment of information exchange and close cooperation between scientific (basic) researchers and investigating officers.

Summarizing, it may be stated that criminal investigations and scientific research share many features but also show fundamental differences. Nevertheless, both disciplines may complement each other. While science can improve established techniques and procedures or develop new ones (among other virtues), the criminal investigator delivers demand and impetus for such advancements. The practical experience of criminalists often initiates new innovations or improvements in techniques. Also, the ultimate test for any (academic) theory is (forensic) practice. Successful implementation and general recognition of a technique is therefore an acknowledgement of quality – unless a refined process is employed – thus analogous to the peer review system. Consequently, a close collaboration between academic research and practical demand is essential.

In addition, both theoretical knowledge and practical experience are indispensable for the education and training of scientists as well as officers working in this field. The crucial position of the expert in the legal system requires a sufficient number of undisputed experts for the wide field of forensic questions. It is generally known that the chances of solving a case decrease drastically after about 48 hours. It is therefore desirable to have highly qualified investigators at a crime scene as soon as possible.

A strong empirical component is certainly present (and essential) in forensic work. Nevertheless, a sound basis in 'framework' sciences like statistics, physics, chemistry, psychology, etc. – is required before specializing in a particular field. Several proposals aim at improving the qualification of examiners (e.g. [27]). However, the need for improvement has also been emphasized for the scientific basis of forensic approaches (e.g. [28]). Therefore, every initiative to advance scientific standards in forensic investigations must be appreciated. An intense and continuous interchange between academic research and forensic-criminalist practice seems the most advantageous strategy to ensure optimum forensic performance.

References

1] European Convention on Human Rights. (1950) European Court of Human Rights, Council of Europe, Strasbourg. 1950

2] K. Popper, Objektive Erkenntnis. Ein evolutionärer Entwurf. Hoffmann & Campe, Hamburg, 1993.

3] S.R. Gross, M. Schaffner, Exonerations in the United States, 1989 –2012, Report by the National Registry of Exonerations, University of Michigan Public Law Working Paper 277, 2012, pp. 1–103.

4] W.C. Thompson, (2009) Painting the target around the matching profile: the Texas sharpshooter fallacy in forensic DNA interpretation. Law, Probability and Risk 8 (2009) 257–276.

5] Frye vs.United States. D.C. 293 F. 1013 (1923).

6] Daubert vs. Merrell Dow Pharmaceuticals, 509 U.S. 579. (1993) p92-102.
7] M. Du, Analysis of errors in forensic science. J Forensic Sci Med 3 (2017) 139-43.
8] J. Ioannidis, Why most published research findings are false. PLoS Medicine 2 (2005) 697-701.
9] A. Sokal, Transgressing the boundaries: Towards a transformative hermeneutics of quantum gravity. Social Text 46/47 (1996) 217-252.
10] J. Lindsay, P. Boyle, The conceptual penis as a social construct. Cogent Social Sciences 3 (2017) 1-7.
11] I. Dror, R. Morgan, C. Rando, S. Nakhaeizadeh, The bias snowball and the bias cascade effects: Two distinct biases that may impact forensic decision making. J. Forensic Sci. 62 (2017) 832-833.
12] H. Ditrich, Cognitive fallacies and criminal investigations. Science and Justice 55 (2015) 155-159.
13] R. Huff, A. Rattner, E. Sagarin, Guilty until proved innocent: Wrongful conviction and public policy. Crime and Delinquency 32 (1986) 518-544.
14] D Canter, L. Hammond, D. Youngs, Cognitive bias in line-up identifications: The impact of administrator knowledge. Science and Justice 53 (2013) 83-88.
15] B. Kamins, ed., Preliminary Report of the New York State Bar Association's Task Force on Wrongful Convictions. New York Bar Foundation, New York, 2009.
16] H. Katterwe, G. Brandes, R. Eisgruber, W. Grimmer, W. Küppers, W. Marquard, K.H. Pohl, Harmonisierte Befundbewertungsskala für kriminaltechnische Untersuchungen. Kriminalistik 12/2007 (2007) 745-750.
17] S. Bunch, G. Wevers, Application of likelihood ratios for firearm and toolmark analysis. Science and Justice 53 (2013) 223-229.
18] C. Mullen, D. Spence, L. Moxey, A. Jamieson, Perception problems of the verbal scale. Science and Justice 54 (2014) 154-158.
19] R. Marquis, A. Biedermann, L. Cadola, C. Champod, L. Gueissaz, G. Massonnet, W.D. Mazzella, F. Taroni, T. Hicks, Discussion on how to implement a verbal scale in a forensic laboratory: Benefits, pitfalls and suggestions to avoid misunderstandings. Science and Justice 56 (2016) 364-370.
20] C. Roux, B. Talbot-Wright, J. Robertson, F. Crispino, O. Ribaux, The end of the (forensic science) world as we know it? The example of trace evidence. Phil. Trans. R. Soc. B 370 (2015) 1-8.
21] R. Julian, S.F. Kelty, Forensic science as "risky business": identifying key risk factors in the forensic process from crime scene to court. Journal of Criminological Research, Policy and Practice 1 (2015) 195 – 206.
22] T.A. Busey, G. R. Loftus. Cognitive science and the law. Trends in Cognitive Sciences 11 (2007) 111-117.
23] I. Dror, D. Charlton, Why experts make errors. Journal of Forensic Identification 56 (2006) 600-616.
24] P.A. Fraser-Mackenzie, I. Dror, Selective information sampling: Cognitive coherence in evaluation of a novel item. Judgment and Decision Making 4 (2009) 307-316.
25] G. Edmont, D. Hamer, A little ignorance is a dangerous thing: engaging with exogenous knowledge not adduced by parties. Griffith Law Review 25 (2016) 282-413.
26] I. Dror, Cognitive neuroscience in forensic science: understanding and utilizing the human element. Phil. Trans. R. Soc. B 370 (2015) 1-8.
27] G. Sapir, Qualifying the expert witness: A practical Voir dire. Forensic Mag. 2/3 (2007). https://www.forensicmag.com/article/2007/01/qualifying-expert-witness-practical-voir-dire (access 7.11. 2017).
28] National Research Council, Strengthening forensic science in the United States: A path forward. National Academies Press, Washington. 2009. http://www.nap.edu/catalog/12589.html (access 21. 4. 2011).