



# Commentary: A “Source” of Error: Computer Code, Criminal Defendants, and the Constitution

Duncan A. Taylor<sup>1,2\*</sup>, Jo-Anne Bright<sup>3</sup> and John Buckleton<sup>3</sup>

<sup>1</sup> Forensic Science South Australia, Adelaide, SA, Australia, <sup>2</sup> School of Biological Sciences, Flinders University, Adelaide, SA, Australia, <sup>3</sup> Institute of Environmental Science and Research, Auckland, New Zealand

**Keywords:** source code, STRmix, DNA profile interpretation, closed source, court challenge

A commentary on

**A “Source” of Error: Computer Code, Criminal Defendants, and the Constitution**

by Chessman, C. (2017). *Calif. Law Rev.* 105, 101–193.

Chessman (2017) warns of the current trend to admit into court unchallenged the results of complex computerized calculations. He provides a number of examples and arguments claimed to demonstrate the need for open source software to remove the “black box” element. We agree with parts of this sentiment, and the topic of this special issue, that there is a danger with those using and receiving information from black box systems.

Some care however is needed with simple diagnoses and prescriptions such as these.

Modern probabilistic genotyping software are replacing methods previously applied manually. We have great confidence in the forensic community with regard to both integrity and dedication. The previously applied processes are usually a composite of standard operating procedure and human judgment. The difference between these and probabilistic software is largely that the processes in the software are encoded.

Many disciplines are sufficiently broad that practitioners need to rely, in part, on the work of others. This is not new (for a discussion on this point see Taylor, 2016). The risk to which Chessman refers arises when the individual using the system has so little understanding that they do not know how to use the system, or when it has not worked<sup>1</sup>. Chessman provides some helpful suggestions for how breaking down black box barriers can be addressed on an individual and systemic scale. As developers of expert system STRmix<sup>TM2</sup> (Taylor et al., 2013), we wish to address some of the alarmist points in Chessman (and echoed by others<sup>3</sup>) that gives the impression that producers of expert systems are all either incompetent or corrupt.

We first wish to correct a couple of points in (Chessman, 2017). Regarding the “erroneous assumption” referenced by footnotes 49–51: This miscode, and indeed any miscode found that has been identified in STRmix<sup>TM</sup> development or use, was identified by examination of the program’s output and not the source code. It would be nearly impossible to identify subtle errors in code by viewing the code. The identification has always been a result of comparison of the results produced by a program to some known control<sup>4</sup>. The results of these comparisons then trigger the examination of a specific section of the code in order to discover the source of the discrepancy.

## OPEN ACCESS

### Edited by:

Sue Pope,  
Principal Forensic Services, UK

### Reviewed by:

Peter Ronald Gunn,  
University of Technology Sydney,  
Australia

### \*Correspondence:

Duncan A. Taylor  
Duncan.Taylor@sa.gov.au

### Specialty section:

This article was submitted to  
Statistical Genetics and Methodology,  
a section of the journal  
Frontiers in Genetics

**Received:** 03 November 2016

**Accepted:** 28 February 2017

**Published:** 15 March 2017

### Citation:

Taylor DA, Bright J-A and Buckleton J  
(2017) Commentary: A “Source” of  
Error: Computer Code, Criminal  
Defendants, and the Constitution.  
*Front. Genet.* 8:33.  
doi: 10.3389/fgene.2017.00033

<sup>1</sup>Note that this is not an issue with just computer programs, recent history has numerous examples within forensic biology showing that a misunderstanding of the way a system works at a fundamental level can cause issues even when the calculations themselves are relatively simple and able to be done by hand (Budowle and Bieber, 2015).

<sup>2</sup>An expert system that analyses STR DNA profile data.

<sup>3</sup>For example see EPIC (<https://epic.org/state-policy/foia/dna-software/>).

<sup>4</sup>Commonly a “by-hand” recreation of the expected value(s).

Even as developers, during the developmental validation of new versions of STRmix™, we utilize the extended outputs of the software to validate, and do not validate by examination of code. A further reference (footnote 98) makes the same incorrect assumption that it was code review that led to the discovery of a programming error. Our experience has been that even more crucial than a review of source code, is the ability to have access to outputs that demonstrate each step of a calculation. We should also note that our ongoing evaluation and testing of the software is a marker of continuous validation and refinement, rather than just fixing “errors” and “blunders.”

The second point we wish to make is that the type and magnitude of miscodes are important to consider. The majority of programming errors will lead to instances of a program “crashing” or failing to produce an answer. These types of errors are arguably inconsequential as they will not lead to any erroneous results being produced. More serious are miscodes where no errors are identified or displayed by the software. These can be split into those that will be clearly identifiable<sup>5</sup> and those that are more subtle and may go initially unnoticed. Even in this latter category, the question should be asked “What effect does this error have?” If the magnitude of the difference in the result caused by the miscode is small compared with the natural variability in the results being produced<sup>6</sup> then arguably the consequences are minimal. We are by no means suggesting that these types of errors are acceptable, they should be rectified as soon as found. We simply suggest that they tend to be used for scaremongering in a manner disproportionate to their impact. Case in point is the oft quoted article (David Murray, 2015), which contains the never quoted sentence “*The DNA likelihood ratios in both the new and original statements appear to be the same.*”

We agree with the suggestion of Chessman that source code should be available for scrutiny. STRmix™ abides by one of the mechanisms that Chessman suggests, namely the ability for code to be disclosed under confidentiality agreements<sup>7</sup>. We note that running of STRmix™ is just the final step in a long journey of computerized activities that ultimately lead to an answer.

<sup>5</sup>Such as value of a probability greater than one, or a negative amount of some substance.

<sup>6</sup>This may either be in the raw results due to inherent variability in the laboratory process or it may be variability in the statistical result due to an evaluation method that utilizes random number generation (Bright et al., 2015).

<sup>7</sup>The code of STRmix™ has been viewed under such conditions in the past.

## REFERENCES

- Bright, J.-A., Stevenson, K. E., Curran, J. M., and Buckleton, J. S. (2015). The variability in likelihood ratios due to different mechanisms. *Forensic Sci. Int. Genet.* 14, 187–190. doi: 10.1016/j.fsigen.2014.10.013
- Budowle, B., and Bieber, F. R. (2015). *Final Report on Review of Mixture Interpretation in Selected Casework of the DNA Section of the Forensic Science Laboratory Division, Department of Forensic Sciences, District of Columbia.* Available online at: <http://dfs.dc.gov/page/usao-report-april-2015>.
- Chessman, C. (2017). A “source” of error: computer code, criminal defendants, and the constitution. *Calif. Law Rev.* 105, 101–193.

A true challenge of all steps in the process would require the examination of the source code underlying the Java programming language in which STRmix™ is written, the Windows™ operating system on which it is run, the software used to process the raw electrophoretic data, the software used to collect the raw electrophoretic data from the electrophoresis instrument, the code used to run the electrophoresis instrument, the PCR thermocycler, the quantification instrument and a myriad of no doubt thousands of blocks of code that sit within the numerous Peripheral Interface Controllers that control hardware components.

With the advent of complex computerized evaluation of evidence, there is a shift from a time where an expert can testify to all aspects of the evaluation, to one where, at some level, the workings of an expert system are accepted without absolute understanding. This may initially seem frightening, but an examination of the bigger picture suggests otherwise. It would be difficult to argue that the use of computerized breathalyzers is a backwards step from the reliability of the Field Sobriety Test. Similarly, virtually all senior advisory bodies relating to DNA profile evaluation recognize the clear benefits of the probabilistic interpretation systems (which by nature of their complexity require computerized implementation) over the preceding manual or binary interpretation methods (Coble et al., 2015; SWGDAM, 2015). In our efforts to ensure that software is not the “source” of errors, it is important to recognize that even with the noted occurrences of these errors, the current computerized solutions, when used by trained experts, represent a vast improvement to the quality and reliability of evidence presented in court.

## AUTHOR CONTRIBUTIONS

All authors contributed to the discussions and writing of the manuscript. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the author’s organizations.

## FUNDING

Funding to write this manuscript was provided by the author’s institutions only in the sense of allowing work time to be used to develop the document.

- Coble, M. D., Buckleton, J., Butler, J. M., Egeland, T., Fimmers, R., Gill, P., et al. (2015). DNA Commission of the International Society for Forensic Genetics: recommendations on the validation of software programs performing biostatistical calculations for forensic genetics applications. *Forensic Sci. Int. Genet.* 25, 191–197. doi: 10.1016/j.fsigen.2016.09.002
- David Murray (2015). *Queensland Authorities Confirm ‘Miscode’ Affects DNA Evidence in Criminal Cases [Online].* Available online at: <http://www.news.com.au/national/queensland/queensland-authorities-confirm-miscode-affects-dna-evidence-in-criminal-cases/news-story/833c580d3f1c59039efd1a2ef55af92b> [Accessed].
- Scientific Working Group on DNA Analysis Methods (SWGDAM). (2015). *Guidelines for the Validation of*

*Probabilistic Genotyping Systems [Online]*. Available online at: [http://media.wix.com/ugd/4344b0\\_22776006b67c4a32a5ffc04fe3b56515.pdf](http://media.wix.com/ugd/4344b0_22776006b67c4a32a5ffc04fe3b56515.pdf) [Accessed 3 October 2016].

Taylor, D. (2016). Is technology the death of expertise? *Forensic Sci. Int. Genet.* 24, e1–e3. doi: 10.1016/j.fsigen.2016.06.006

Taylor, D., Bright, J.-A., and Buckleton, J. (2013). The interpretation of single source and mixed DNA profiles. *Forensic Sci. Int. Genet.* 7, 516–528. doi: 10.1016/j.fsigen.2013.05.011

**Conflict of Interest Statement:** Authors are technical developers of commercial software STRmix™ but do not benefit financially from STRmix™.

Copyright © 2017 Taylor, Bright and Buckleton. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.