Digitising History From Below: The Old Bailey Proceedings Online, 1674-1834

Professor Tim Hitchcock, University of Hertfordshire
Professor Robert Shoemaker, University of Sheffield

Abstract
   The Old Bailey Proceedings Online has made available in a fully searchable online edition the largest body of texts detailing the lives of non-elite people ever published. This article explains the origins, methodologies, and outcomes of this project, and assesses how the website has been used in academic teaching and research since its launch in 2003. The limitations as well as the benefits of providing access to primary research materials in this medium are considered. It concludes with an outline of current plans for the digitisation, and integration into the current website, of further substantial bodies of digitised sources on related topics.

   Between 1674 and 1834 the Proceedings of the Old Bailey were published eight times a year, providing accounts of all the trials for serious crimes committed in London and Middlesex. First produced as part of a wider explosion of printed literature about crime in the late seventeenth century, they fed an immense popular interest in crimes and criminal lives. The Proceedings quickly become an established periodical and were avidly read by Londoners eager to hear about the latest gruesome murder or violent robbery. Over 160 years they detail 100,000 trials, and include over 60,000 pages of text, representing the largest single source of information about non-elite lives ever published. They provide a wealth of detail not only about crimes, the judicial process, and punishments, but also about everyday life in preindustrial London. They describe life on the streets and in pubs, coffee houses, workplaces, and lodgings; and they open a window onto the experiences of Londoners of all ages, classes, and backgrounds. The Proceedings are now accessible throughout the world at the click of a mouse owing to the creation of a website which, since its launch in 2003, has been widely used in teaching and research.

The Project
   Previously the twenty-five million words contained in these publications could only be accessed easily by reading, in date order, either the over one thousand published editions in hard copy or a collection of 38 reels of microfilm.[1] Anyone wanting to search the Proceedings for a particular type of information, a place or type of crime for instance, had to devote months to the task.   Those who went to the trouble produced some important works, notably Dorothy George's seminal London Life in the Eighteenth Century, John Langbein's magisterial study of the introduction of lawyers in criminal trials, Hans-Joachim Voth's analysis of the impact of the industrial revolution on working hours, and Dana Rabin's work on the evolving language of the self and ideas of legal responsibility in the eighteenth century.[2] But the potential this resource possessed to shed light on a vast array of other research topics remained underexploited owing to the sheer extent of the text involved. The creation of an electronically searchable digital edition has allowed researchers to find what they want in seconds, but it was an expensive undertaking. Two generous grants (totalling £700,000) from the Arts and Humanities Research Board (now

Arts and Humanities Research Council—AHRC) and the New Opportunities Fund (a lottery distributor), allowed the Proceedings to be digitised and posted on the internet, making them available free of charge to web users throughout the world.

The project created digital images of both the original pages and their textual content. Because the original typeface is so irregular, it was not possible to use an optical character reader, and the text had to be 'double rekeyed': that is, typed in twice and the results electronically compared in order to identify errors. This mammoth task was managed by the Higher Education Digitisation Service at the University of Hertfordshire. The text was then transferred to the Humanities Research Institute at the University of Sheffield, where it was prepared for web searching. This involved, first, inserting invisible XML tags into the text delineating the different types of information that can be searched, including names, crimes, verdicts, punishments, crime locations, lists of juries, and advertisements. Then a search engine had to be developed to facilitate web-based searching. This was no trivial task, since XML search engines were then (2002-3) at a very early stage of development, and the Proceedings constituted one of the largest XML databases in existence. A customised search engine was constructed, based on a piece of shareware called eXist (later replaced by Lucene), and then spliced to an existing database programme (MySQL). Finally, a set of web pages was written explaining the background and context of the Proceedings for users unfamiliar with the source.

Results

The Old Bailey Proceedings Online (www.oldbaileyonline.org) launched in March 2003 and completed in July 2005, can be searched in a number of different ways. Most importantly, users can search the 25 million words for any text string, such as 'coffee house', 'handkerchief', 'carpenter', or 'boxing match'. (Wild cards can be used to allow for variant spellings). Structured searching of text marked up with XML tags allows one to search by name, type of crime, type of punishment, verdict, crime location, crime date, defendant and victim occupations, and age of defendant (if convicted). An innovative statistical search page allows the creation of pie charts, bar charts, and tables tabulating such variables as crime type, verdict, punishment, and defendant or victim gender. It is now the work of a few moments to tabulate the rate of prosecution of a particular crime over 160 years, or the changing gender balance among those brought to trial. [screen shot of table showing defendant gender by decade]

Two other search pages allow users to place Old Bailey evidence in the context of other sources. A place and map search allows users to see the locations of crimes and defendant homes plotted on digitised contemporary maps of London, and a linked associated records database lists other manuscript and printed materials relating to specific trials, providing researchers with the information necessary to continue their research in libraries and archives in the UK. In addition, digitised images and texts of the some of the associated records, notably the 'Ordinary's Accounts' (biographies of executed criminals written by the chaplain of Newgate Prison) and the manuscript sessions papers (including witness depositions and examinations of suspects), have been incorporated for the period 1745-55. This allows researchers to examine several related accounts of the same crime in order to gain a fuller picture of the events which led to a trial at the Old Bailey.

Usage of the website has exceeded our wildest expectations. In its first two years it recorded more than a million visits and the site now clocks up over 6,000 visits a day from users throughout the world. By far the most popular search page is the name search, reflecting huge interest from genealogists, but there is also substantial usage of all the other search and historical background pages. The site is widely used in undergraduate and postgraduate teaching, and by researchers in several academic disciplines.

Teaching

Many of the 126,000 WebPages in English which currently include the website URL are from schools, colleges and universities, which show that the site is widely used as a resource for teaching. A section of the website, 'For Schools', provides suggestions for how the resource can be used in teaching students in the UK's Key Stages 2 through 5 (ages 9+) for developing skills not only in historical enquiry and information technology, but also in literacy, numeracy, citizenship and cultural diversity.[3] Pages in this section provide advice for Teachers, selected primary sources and images, a timeline and glossary, the story of the journey from Newgate to Tyburn, and a number of suggested 'tasks for students'.

The site is widely used in undergraduate and postgraduate history teaching throughout the English speaking world. Entire documents-based modules can be constructed around it, such as the University of Sheffield's final year special subject, 'Crime and the Law in Eighteenth-Century London'.[4] Individual seminars within courses can use the site for key reading, as in a seminar on 'Gay subcultures ca 1710-1850' in the Birkbeck College continuing education module on 'Queer Histories: London's Lesbian and Gay Past'. In preparation for this seminar, students were asked to search the site using the keywords 'sodomy', 'molly', and 'buggery'.[5] Similarly, students taking a seminar in the Kings College London MA module 'The Body and Society in Early Modern Europe' on the subject of 'Diseased Bodies' were asked to search the site for the word 'pox'.[6] The website can also be used for skills teaching. At the University of Glasgow, the website is used in seminar teaching to teach source criticism and explore key themes on the second year module 'Government, Culture and Society in Early Modern Europe, 1550-1715'. Students are asked to read some of the historical background pages, search the proceedings for trials of interest, and perform statistical searches, with one seminar focusing on verdicts and punishments. They can also choose to write a brief seminar paper involving the analysis of up to three trials and a statistical search.[7] The resource is also used in more substantial written assignments: a course in 2003 at the University of Calgary on 'Crime and Criminal Justice in England' assigned a research essay which asked students to search the site to identify between 10 and 20 trials of interest. The students were then instructed to reflect on them in light of the course themes and situate their evidence within the context of the relevant secondary literature.[8] It would also be possible to use the website for exercises involving statistical analysis and quantitative reasoning.

By far the most common use of the website in undergraduate and postgraduate teaching, however, appears to be as a resource for final year undergraduate or MA dissertations and research projects. A vast array of potential topics in the history of crime, justice, and daily life in London can be researched using the Proceedings. The availability of complementary electronic resources such as the printed pamphlets in

Eighteenth-Century Collections Online and newspaper reports in the Times Digital Archive means that students can contextualise their findings by consulting a wide range of additional sources.

Pages recently introduced on the website give teachers, lecturers and researchers the opportunity to report back on their use of the resource, and it is hoped that these pages will facilitate a fruitful exchange of ideas between website users.

Research Impact

The manifold research possibilities opened up by this resource mean that it has the potential to transform the way historical research is conducted. By making such a wealth of material about the lives of ordinary people accessible, it facilitates the creation of a truly democratic history, both in the sense that this resource is available to anyone and because it is about the lives of everyone who lived in eighteenth-century London. More than this, the publication of a resource which can concurrently be both textually searched and statistically analysed allows quantitative and qualitative research to be joined, thereby enhancing the authority and relevance of both.

An international conference held at the University of Hertfordshire in July 2004 demonstrated some of the research topics which have been opened up by the digitised resource. The ability to search millions of words of text to locate details of specific topics, underpinned papers on subjects as diverse as bigamy, homosexuality, soldiers' and sailors' marriages, the Irish poor in London, urban fairs, the London privy, and female murderesses. Some of the papers were subsequently published in a special issue of the London Journal.[9]

First and foremost, the resource provides a rich resource for historians of crime and criminal justice, contributing to an already extensive historiography.[10] The statistics function allows historians to obtain more precise evidence of phenomena previously only partially documented, such as the surge in prosecutions of women for theft in the late seventeenth century (to the point where they constituted over half of the defendants prosecuted, a rare statistic in the history of crime), or the impact of transportation in 1718 on sentencing patterns for felons.[11] But as the research interests of historians of crime turn to language, ideas, and mentalités,[12] it is the textual narratives which are likely to be most intensively consulted. Not only can historians assess the evolving meanings of terms such as 'the mob', 'thief-taker', 'cunning man' and 'police', but they can analyse the narrative strategies used by the poor in courtroom testimonies, identifying some of the ways in which they attempted to manipulate the judicial system for their own ends. Current work is combining research on the Proceedings with other genres of printed literature about crime, including the Ordinary's Accounts, newspapers, broadsides and ballads, and even the novel in order to study the emergence of modern attitudes and subjectivities.[13]

Indeed, the often evocative language of the Proceedings provides rich material for scholars in literature and other disciplines. For example, Professor Lincoln Faller at the University of Michigan has analysed rape cases in the context of his study of Samuel Richardson's Clarissa, arguing for reciprocal influences between the often sentimental language of printed trial reports and the novel.[14] Linguistics experts are mining the Proceedings for evidence of changes in spoken language at the time of the transition from early modern to modern English: Dr Magnus Huber of the University of Giessen

(Germany), for example, is creating a Linguistic Corpus of spoken English from the Proceedings, and analysing differences according to the social parameters of age, gender, place of origin, and social status.

One of the biggest advantages of digitisation is that resources like this can be used by scholars working on a wide range of topics who would not otherwise have had the time to read through the Proceedings. Many research projects will use this resource as a supplement to the main body of sources consulted. Thus, Tim Hitchcock's work on the poor has been radically transformed by the ability to search for a single place or name or thing or action. In particular, by tracing individual paupers found in a range of archives and records series also on the website, he has been able to place the poor's experience of the complex landscape of social policy provision at the centre of the analysis. In the process a new series of questions about the impact of pauper agency on the evolution of hospitals, prisons and workhouses have been raised.[15]

Some of the more innovative aspects of the site have yet to be systematically exploited. The possibilities for studying the geographical distribution of criminal prosecutions opened up by the mapping feature should lead to new understandings of London's notorious criminal hotspots such as Drury Lane and Black Boy Alley. The tagging of evidence of the ages of those convicted after 1788 allows study of the criminal lives of different age groups, and their differential treatment by the courts. And comparison of the representations of crimes in the Proceedings with the manuscript depositions concerning the same crimes which were created before the trials and the Ordinary's Accounts and other related publications which followed executions will allow historians to study how criminal narratives were selectively reinterpreted when they were represented in print.[16] This has prompted Robert Shoemaker to investigate the vital new role played by print in the creation of public knowledge about crime in the eighteenth century.

Drawbacks

Is there a downside to the creation of a resource like this? Like any body of evidence, the Proceedings must be used with care and with recognition that it is only one of a number of possible sources for the study of most topics. But there are particular dangers associated with digitised resources provided on the internet. There will, of course, be some transcription errors, and even errors in the XML tagging, but the provision of digital images of the original pages allows users to check the transcription when in doubt, and periodic updating of the website will allow transcription and tagging errors to be corrected. More serious is the danger that users, perhaps directed to the site by a Google search, will hastily extract information without a proper understanding of the nature of the source and the context of the particular trials examined, and without consulting relevant non-digitised sources.[17] Here it is the duty of users to conduct their research properly, but even those familiar with the source experience it on the internet differently from eighteenth-century or even twenty-first century readers who read the Proceedings in hard copy as a periodical, not as a vast database from which relevant information can be extracted from texts published across a long time period. It is important to experience the document in its original form, and users can do this by using the 'Browse by Date' search page to select editions of the Proceedings which can be read in their entirety, rather than broken up into individual trials. Finally, the accessibility and prominence of a website

like this will inevitably draw more attention to this source, at the expense of possibly richer sources which are not so easily accessible. Here the 'Associated Records' feature of the site is intended to remind readers that other sources for the study of many crimes are available in libraries and archives. In the longer term, the ongoing digitisation of other sources will make many other primary sources equally available.

Future Plans
    The Old Bailey Proceedings did not cease publication in 1834; they were simply given a new name to reflect the court's new title and jurisdiction as the Central Criminal Court, and publication continued until it suddenly and permanently ceased in 1913. Funding from the AHRC has been awarded to allow the 100,000 trials from this period to be added to the website by 2008, enabling users to conduct searches over the entire 240 year period from 1674 to 1913, from just after the Great Fire to just before the Great War.
    In the longer term, by using the power of information technology and XML markup to bring together a wide range of texts and records, an expanded version of the Old Bailey Proceedings Online will allow the creation for the first time of a fully rounded picture of daily life in the 'long' eighteenth century. A project funded by the Economic and Social Research Council, 'Plebeian Lives and the Making of Modern London, 1690-1834' will digitise a significant proportion of administrative records from eighteenth-century London relating to the provision of social welfare. Specially developed record-linkage programmes will integrate these records with the rich array of already existing electronic databases of eighteenth-century sources, including the Old Bailey Proceedings Online, in order to create a composite picture of individual interactions with the agencies of charity and local and central government. This will allow historians to examine the interrelationships between the histories of criminal justice, poor relief and medical care, and ascertain how individual Londoners used and experienced all these agencies in their day-to-day lives, refocusing the historiography of social welfare from institutions to individual experiences.[18] The ultimate aim of this project is to study how patterns of public use of justice, poor relief, and medicine shaped the evolution of the modern welfare state. Judging by the experience of the Old Bailey Proceedings Online, however, the resulting web-based resource will also be used to research a myriad of other research questions we cannot even anticipate, and the resource will continue to grow as it is linked with other newly digitised sources.

Bibliography

Beattie, J. M., *Crime and the Courts in England 1660-1800* (Princeton: Princeton University Press, 1986).
Beattie, J. M., *Policing and Punishment in London 1660-1750* (Oxford: Oxford University Press, 2001).
Gaskill, M., *Crime and Mentalities in Early Modern England* (Cambridge: Cambridge University Press, 2000).
George, M. D., *London Life in the Eighteenth Century*, 2nd ed. (London: Penguin Books, 1966).
Gladfelder, H., *Criminality and Narrative in Eighteenth-Century England* (Baltimore and London: Johns Hopkins University Press, 2001).

Hitchcock, T., *Down and Out in Eighteenth-Century London* (London: Hambledon and London, 2004).

Hitchcock, T. and Shoemaker, R., eds., 'Special Issue: Tales from the Old Bailey', *London Journal*, 30/1 (2005).

King, P., *Crime, Justice and Discretion in England 1740-1820* (Oxford: Oxford University Press, 2000).

Langbein, J. H., *The Origins of Adversary Criminal Trial* (Oxford: Oxford University Press, 2003).

McKenzie, A., 'From True Confessions to True Reporting? The Decline and Fall of the Ordinary's Account', *London Journal*, 30/1 (2005): 55-70.

Palk, D., 'Conference Report', *London Journal*, 29/2 (2004): 54-56.

Poole, S., 'Report Back', *History Workshop Journal*, 59 (2005): 282-84.

Rabin, D., *Identity, Crime, and Legal Responsibility in Eighteenth-Century England* (Basingstoke: Palgrave Macmillan, 2004).

Shoemaker, R., 'Digital London: Creating a Searchable Web of Interlinked Sources on Eighteenth-Century London', *Program: Electronic Library and Information Systems*, 39/4 (2005): 297-311.

Shoemaker, R., *The London Mob: Violence and Disorder in Eighteenth-Century London* (London: Hambledon and London, 2004).

Voth, H-J., *Time and Work in England 1750-1830* (Oxford: Clarendon Press, 2000).

Walker, G., *Crime, Gender and Social Order in Early Modern England* (Cambridge: Cambridge University Press, 2003).

For further reading, consult the Old Bailey Proceedings Online Bibliography.

---

[1] The collection, published by Harvester Microform, covered the period from 1714 to 1834. Copies of the Proceedings published before 1714 were kept in a number of disparate libraries and had never before been assembled in a single collection.

[2] M. D. George, *London Life in the Eighteenth Century* (1925; 2nd edn, London: Penguin Books, 1966); J. H. Langbein, *The Origins of Adversary Criminal Trial* (Oxford: Oxford University Press, 2003); Hans-Joachim Voth, *Time and Work in England 1750-1830* (Oxford: Clarendon Press, 2000); D. Rabin, *Identity, Crime, and Legal Responsibility in Eighteenth-Century England* (Basingstoke: Palgrave Macmillan, 2004).

[3] http://www.oldbaileyonline.org/schools.

[4] http://www.shef.ac.uk/history/current_students/undergraduate/modules/level_3/hst3019-20.html.

[5] http://www.bbk.ac.uk/study/ce/pdfs/hi/HI144UACB.

[6] http://www.kcl.ac.uk/humanities/history/courses/body%20syllabus%202005.doc.

[7] http://www.arts.gla.ac.uk/History/Modern/Level2/index.htm#info.

[8] http://hist.ucalgary.ca/courses/sp2003/HTST473L20.htm.

[9] 'Special Issue: Tales from the Old Bailey', *London Journal*, 30/1 (2005). For reports of the conference, see Deirdre Palk, 'Conference Report', *London Journal*, 29/2 (2004): 54-56; S. Poole, 'Report Back', *History Workshop Journal*, 59 (2005): 282-84.

[10] See, in addition to the works cited elsewhere in these notes, P. King, *Crime, Justice and Discretion in England 1740-1820* (Oxford: Oxford University Press, 2000); R. Shoemaker, *The London Mob: Violence and Disorder in Eighteenth-Century London* (London: Hambledon and London, 2004); G. Walker, *Crime, Gender and Social Order in Early Modern England* (Cambridge: Cambridge University Press, 2003).

[11] J. M. Beattie, *Crime and the Courts in England 1660-1800* (Princeton: Princeton University Press, 1986), chaps. 9-10; J. M. Beattie, *Policing and Punishment in London 1660-1750* (Oxford: Oxford University Press, 2001), 63-71, 432-62.

[12] M. Gaskill, *Crime and Mentalities in Early Modern England* (Cambridge: Cambridge University Press, 2000).

[13] See A. McKenzie, 'From True Confessions to True Reporting? The Decline and Fall of the Ordinary's Account', *London Journal,* 30/1 (2005): 55-70; H. Gladfelder, *Criminality and Narrative in Eighteenth-Century England* (Baltimore and London: Johns Hopkins University Press, 2001).

[14] Lincoln Faller, 'The *Clarissa* Effect'. We are grateful to Professor Faller for showing us a copy of his (as yet) unpublished essay.

[15] T. Hitchcock, *Down and Out in Eighteenth-Century London* (London: Hambledon and London, 2004).

[16] For some preliminary findings comparing representations of the same crime in different sources, see 'The Value of the Proceedings as a Historical Source' on the Old Bailey website: http://www.oldbaileyonline.org/proceedings/value.html.

[17] Poole, p. 282.

[18] R. Shoemaker, 'Digital London: Creating a Searchable Web of Interlinked Sources on Eighteenth-Century London', *Program: Electronic Library and Information Systems*, 39/4 (2005).