

Evidence-Based Genealogy

by Mark B. Arslan

e-mail: contact@arslanmb.org

8 June 2024

Abstract

Many of the published family trees, especially those posted online, are poorly documented with respect to primary documentary sources. As a result, it is quite difficult to confirm the reliability of the information presented – names, relationships, dates, and places. Most genealogy software programs allow the exporting and importing of GEDCOM files, which is a mixed blessing. It is now much easier to share your findings, but errors due to misinterpretation of source material, or just plain shoddy research, can quickly be replicated a thousand-fold and work its way into online genealogy databases. (When you see something restated enough times, it takes on a life of its own and becomes “gospel”.) The good news is that technology has also greatly improved the availability of primary source material, much of it now easily accessible online. By making extensive use of primary documentary sources in your own research, augmented with findings from DNA testing, you can create a clearly documented evidence trail to improve the quality of your own work and more accurately assess the reliability of the work of others as you decide what you want to incorporate into your own family trees.

Background

As a young boy, I was very close to my maternal grandfather. I often enjoyed hearing him talk about growing up in rural Arkansas, surrounded by a multi-generational extended family. I was fascinated with trying to understand the relationships between the different characters in his stories and started creating charts to keep them straight. At age 11 (in 1968), that led to a life-long obsession with genealogy.

Grandpa put me in touch with several older members of his family (1st-cousins and 2nd-cousins) and I started writing to them, sending them questionnaires through the mail, asking about what they knew about their families. They graciously tolerated my questions with prompt responses and put me in touch with others. I'm glad I did that while I was young, as they are all long gone. I quickly realized that much of the information I gathered from the questionnaires was incomplete and one person's recollections didn't always agree with another person's. Also, the information I gathered from the questionnaires only reached back a few generations. How would I resolve the discrepancies, fill in the gaps, and extend my knowledge further back?

In those days, access to primary documentary sources was very limited, compared to today. You could order microfilms of the US census records up to the year 1880 from

the National Archives and Records Administration (NARA) (as well as other federal documents like military service and pension files), visit courthouses to look at the original records, or find abstracts (secondary sources) in the genealogy section of local libraries. That was about it. When I was 16 years old (1973), Grandpa took me and a cousin across country to his hometown of Ozark, Arkansas for a wonderful visit and I got to spend an entire day digging through the county courthouse there and visiting old cemeteries.



Franklin County, Arkansas Courthouse (Ozark, 1973)

Early on I made a few decisions that paid dividends later. First, when I reached out to older relatives, I asked many of the same questions to each of them, to try to see if I'd get the same answers (corroborate the information). Second, I learned to carefully document my sources of the information as I built my family trees – for dates and places of birth, marriage, and death, as well as filiation. (Those source citations, numbering in the tens of thousands, proved immensely helpful decades later as many of the primary sources became readily available online.) Third, I learned how to program computers as a teenager (in the early 1970's) and wrote my own genealogy software to organize my work. (I've enhanced that software many times in the last 50 years and still use it today, now with a MySQL database and php scripts, with decades of abstracts from primary sources captured in Microsoft Word and Excel documents.)

The Problem

As I expanded my work to include more of the families from whom I am descended (a mix of French-Canadian/Acadian, Armenian, German, and British – England, Scotland, Wales, maybe a little of Ireland), I encountered genealogical works on these families published years ago. Some of those genealogies were clearly the result of careful research, with many of the primary documentary sources cited. Others, not so much. I found in many cases that the information presented was in conflict with other works, as well as my own research. It was only by going back to the original primary sources that I was able to resolve the discrepancies and draw my own conclusions. This was much easier when the sources were cited.

A number of genealogical web sites like Ancestry.com and FamilySearch.org now allow contributors to post extensive family trees online. These are often used in conjunction with standalone or web-based genealogical software. There is a standard file format known as [GEDCOM](#) that allows you to save your family trees, move them from one platform to another, and share with others. Citations of sources can be included, as well. Some contributors take full advantage of this capability, carefully documenting primary sources and even including links to images of the sources themselves. (Hats off to them!) Usually, however, the sources are not cited, or the citation is insufficient to allow other researchers to find the source. Too often, what is cited is a tertiary source like a published genealogy or, worse yet, another GEDCOM. For a source to be most useful, it must be easily located by another researcher (preferably online) and it should be as close as possible to a primary source. (More on that later.)

Many contributors of family trees (both ancestral and descendant) to genealogy platforms opt for quantity over quality. Their goal, apparently, is to add as many ancestors (or descendants) as they can. To accomplish this, they find GEDCOM files that others have posted and then do bulk imports into their own trees, without making any effort to first validate the information. This often results in family groups that make no sense – duplication of people, ages of children whose births occur far outside of the reasonable age for the mother (approximately age 15 to 45), or places of birth, marriage, or death that are nowhere near where the family actually lived. I've seen many examples of information that was debunked decades ago that refuses to "die", given the endless replication by willy-nilly imports of GEDCOMs.

On ancestral trees, the biggest issue is filiation (parent-child relationships) that have no evidence to support them. Newbies to genealogy are often so anxious to extend their lineages far back in time that they don't exercise the proper due diligence to carefully validate each parent-child relationship. It is very common to take an immigrant ancestor and "find" their parent in the Old Country, without any evidence whatsoever. This step (perhaps across the pond from colonial Virginia back to England) requires a high evidentiary standard, such as a Virginia headright grant mentioning a place of origin in England, accompanied by supporting entries in a contemporary English parish register. (Y-chromosome DNA evidence may also be helpful in narrowing down the search to the correct historic English county.) French Canada's very complete parish registers from

the 1600's to 1900's are the best source I've seen for easily proving filiation back 10+ generations, along with the French-Canadian Catholic practice of the woman keeping her maiden name. Parish registers in England, France, Germany, and other European countries are rarely extant prior to the 1500's. Any ancestral tree that purports to trace someone's lineage back earlier than that should be viewed as highly suspect, unless there is compelling evidence from primary documentary sources (which there rarely is). Those that trace back to the Middle Ages (e.g., Charlemagne) or, even worse, to a biblical character, are not to be given any credence.

It is understandable that people was to claim descent from (or relationship to) "someone famous", especially if they share a common surname with one's family. Again, the evidentiary standard for such a claim should be very high, proving filiation each step of the way with primary documentary sources. Just because someone shares a surname doesn't mean they are related to you. (The best way to tell is for men with that surname to do a Y-chromosome DNA test. I manage over a dozen surname-based Y-chromosome DNA projects for FamilyTreeDNA and have found this tool to be extremely helpful in proving, or disproving, relationships to common direct paternal ancestors.)

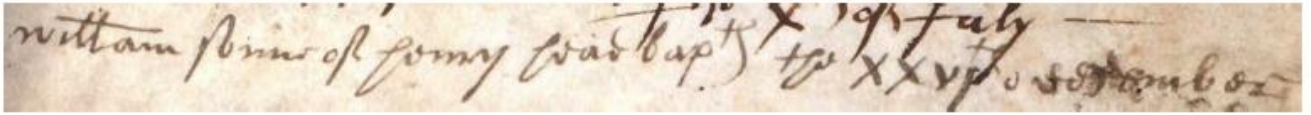
Types of Sources

Genealogical sources can be of several forms. They could be personal remembrances (like an autobiography or an interview), but keep in mind that events that happened many years earlier might be a bit hazy, or the person might not be totally credible regarding some information, for a variety of reasons. Documentary sources that were created at the time of an event (such as a land deed, will, or court minute), many of them hand-written, are considered primary sources and are the most compelling evidence. (Even primary sources can contain inaccurate information, so corroboration with other primary documentary sources is important to seek.)

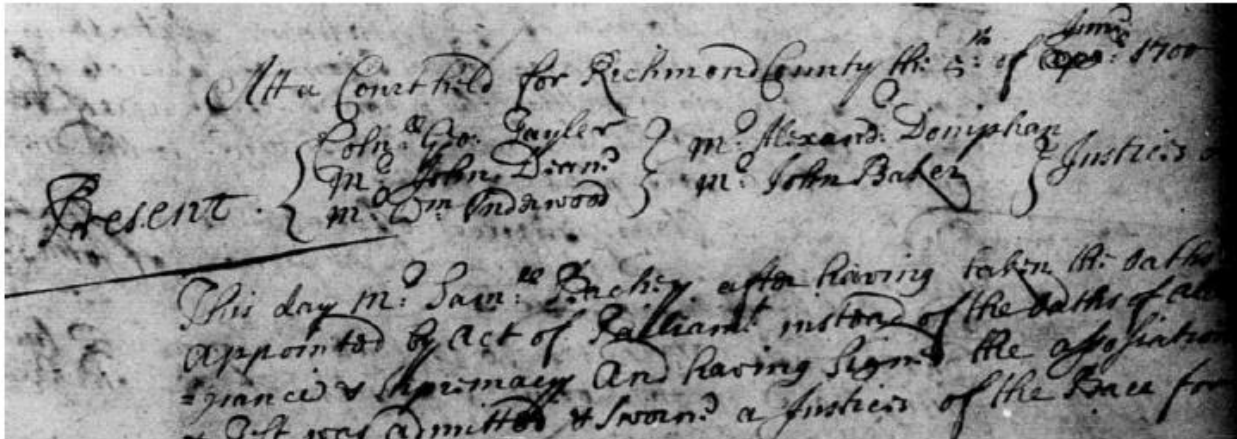
Secondary documentary sources may include a compendium of marriages (from a variety of records like marriage licenses and bonds along with marriage returns and certificates) or something published in a newspaper (like an obituary). Secondary documentary sources are useful, but it is a good idea to validate the information by locating and examining the primary source from which it was derived. Many primary documentary sources prior to the early 1900's were hand-written and the person compiling the secondary source may have not read the handwriting correctly or miscopied it. (Handwriting of earlier colonial America, from the early 1600's to middle 1700's, can be quite challenging to read until you've gotten accustomed to it. That of the late 1500's and early 1600's in England is even tougher. I've seen some secondary sources that transcribed information like names and dates incorrectly.)

26 Dec 1642 Head, William
son of Henry Head

#



Baptism entry (1642) from parish register of Box, Wiltshire County, England



Court minutes entry (1700) from Richmond County, Virginia

Tertiary documentary sources (my term) might include things like county histories, biographies, or published genealogies. They are far removed from the primary sources and their credibility is mostly a function of the quality of the research performed. They should never be taken as wholly factual and should be validated by examining primary documentary sources.

What kinds of primary documentary sources are most helpful? It depends on the location and timeframe. For those whose families came to America fairly recently (like my Armenian relatives who came in the early 1900's from Ottoman Turkey and the Russian Empire), sources like ship manifests, naturalization and passport applications, newspapers, military draft registrations, censuses, and vital records are key. See the article I wrote in The Armenian Weekly magazine:

<https://armenianweekly.com/2021/11/05/armenian-immigration-to-north-america-through-the-1930s-a-compilation-of-primary-sources/>.

When doing genealogical research of the colonial period in America, however, a much different set of resources would need to be consulted. Most helpful would be things like probate records, land and personal property deeds, [court minutes](#), tax lists, church registers, and (if you can find them) Bible records.

The good news is that images of many of these primary documentary sources are available online now at web sites like Ancestry.com, FamilySearch.org, Fold3.com, and others. Some are indexed, many are not. (Indexes are not perfect; sometimes, a page-by-page reading through a handwritten volume may reveal some exciting new

information.) When citing these sources, refer to the original title, volume/book, and page number (along with the date of the entry) (e.g., *VA, Northumberland Co., Court Minutes (1652-1665), p. 130 (5 Sep 1660)*). If you only cite the link (URL) to the entry's image on the web site, that might change and no longer be valid ten years from now.

More recently, DNA test results are a good complement to primary documentary sources, when used properly. Autosomal DNA tests are very useful for corroborating genealogical research to prove relationships more recent than around five generations ago (150 years or so before the age of the person being tested) using triangulation techniques. This assumes that there are relatives who have also been tested at the same company as you (a good bet if they've been in the US for several generations). I highly recommend AncestryDNA due to their large client base, analytical tools like shared matches and custom groups (with colors), but most importantly the relatively high percentage of clients who've submitted ancestral trees. (Be sure to post your own tree, as well.) Many biological parents have been discovered this way.

For more remote ancestry, Y-chromosome DNA testing (through companies like FamilyTreeDNA) is useful. This test is for males only and looks exclusively at the direct male lineage (your father's father's father's father, etc.). Since surnames are often handed down the direct male line, this is a good tool for figuring out if two men with the same surname are actually of the same family. As shown in the project I manage for the Head family (<http://arslanmb.org/yDNA/Head-DNA-yDNAResults.htm>) there are many of the test participants with that surname who do not share the same direct paternal lineage. I've clustered the family groups together by color. For example, the kits grouped together with the bright yellow background all match one another and share virtually the same Y-chromosome DNA. Many of them in the US are descended from the immigrant ancestor William Head, born in 1642 in the parish of Box, Wiltshire County, England, who came to Virginia in 1660. He died around 1710 in what was then Richmond County, Virginia (now King George County). His name and parish of origin (Box) appear on a headright grant. Since the one English participant in the cluster (the kit labeled "Ancestry.com 8") also originated in Wiltshire County, not far from the parish of Box, this evidence corroborates the origins stated in the headright grant. I've had similar success with other surnames like Bradshaw.

Don't assume that information in a primary source is entirely correct. Expect that there may be inconsistencies between primary sources. That's normal. Here are some reasons why. That's why it is important to corroborate by finding as many references as you can for a person in primary sources. It also highlights the importance of citing relevant primary sources in a genealogical work, as other researchers might have a different interpretation of the evidence.

- Informant didn't know
- Lack of understanding between clerk and informant due to language difference
- Inconsistencies in spelling of personal names
- Error in transcribing between work document and official document (or copying of official document)

- Hard to read (poor image or difficult handwriting)
- Places names different as a result of changing borders (e.g., counties of colonial Virginia)
- Delay in time between event and recording of information
- Deliberate falsification (informant or clerk)

Suggestions for Citing Primary Documentary Sources

There are many different ways of citing primary documentary sources in your genealogical work, as much depends upon the platform you are using to store and present your findings, whether it be desktop software, a web site, or just plain paper. Remember, the objective is to allow others to be able to find the evidence that led you to the conclusions that you reached.

When researching primary documentary sources, I carefully abstract each entry, capturing the relevant information and providing enough information about the source to allow me and others to easily locate it later and reexamine it. See <https://markarslan.org/genealogy-public/Head-ColonialVA/Notes-VA.pdf>. Some information, such as dates¹, numbers larger than 10, and currency (£ or \$), I might standardize for readability. For documents like wills, deeds, and tax lists, using a template is helpful, as shown below. When recording information verbatim, enclose it in quotes (“ and ”). Record spelling variations exactly as written in the original document. For readability, it is okay to standardize capitalization and punctuation.

King George Co., Deed Book 2 (1735-1744), pp. 317-319 {34234}

Date written: 27 Jul 1740

Date recorded: 3 Oct 1740

Grantor(s): Alford Head [#11] planter, of Hamilton Parish, Prince William Co.)

Grantee(s): Harry Turner (of Hanover Parish, King George Co.)

Witness(es): Henry Churchman, Thomas Turner, Jr., & William Marshall

Amount paid: 4,000 pounds of crop tobacco

Description: 100 acres of land in Hanover Parish, King George Co. “adjoining to the lands of William Peck, Robert Harrison, Grace Berry, Richard Green, William Duff, & corners on the land whereon William Jameson now dwells belonging to Thomas Turner, which said land was formerly purchased of Robert Peck by William Head the Elder [#] & on his decease the s’d land descended to William Head the Younger [#1], who was heir at law & eldest son of the said William Head the Elder. And, on the decease of the said William Head the Younger, the same became vested in the aforesaid Alford Head, party to these presents, who was the eldest son & heir at law to the said William Head the Younger”

¹ When researching records of colonial America prior to 25 September 1752, be aware of the use of the old Julian calendar, where the new year didn’t start until 25 March. For example, the day after 31 December 1717 would be 1 January 1717. When recording these dates between (and including) 1 January and 24 March, it is helpful to record the year with a slash (/), such as 1 January 1717/8. You’ll often find it shown this way in primary documentary sources of that time, like County Court Minutes in Virginia.

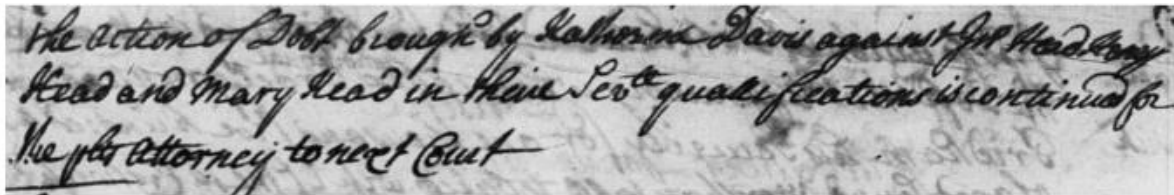
You'll note (above) that I put a number (34234) in braces (curly brackets) (in the red rectangle) after the source citation. This shorthand allows me to easily cite that source later when presenting my findings, such as in a descendants listing. In my abstract, I also provide a descendant number or ancestor number in brackets (in the blue rectangle) that points to that person in my database. When viewing my abstracts, you can easily search for that person by their ID number to find all references to them in that particular file. For my research of records in America, both in the colonial era and after the American Revolution, I arrange my notes by state (one Microsoft Word or Excel document per state) and then by county² within that. I prefer using Excel when recording US federal censuses, and Word for everything else. If the descendant or ancestor number changes, I can easily do a global find/replace in my notes files.

For entries that are of particular importance to your research, consider placing a portion of the original image within your notes (adjacent to your abstract).

Richmond Co., Court Minutes Book 6 (1711-1716)

p. 36 (7 Aug 1712) {34245}

"In the action of debt brought by Katherine Davis against Jno. Head [#4], Henry Head [#5], & Mary Head [#1] in their sev'll qualifications is continued for the plt's Attorney to next Court."



When presenting your findings, you can now easily refer back to the original source citation and your associated abstract. This enhances the credibility of your genealogical work products and allows you and others, many years later, to resume research on those same families.

Here is an example of a listing of descendants, from my research of my Head family of colonial Virginia, along with a description of the format of the listing and the descendant ID. (This same information is available in a GEDCOM format for importing the tree into a website like Ancestry.com.)

² Make yourself familiar with the frequent formation of new counties, as well as county shape and border changes, for the area in which you are doing your research. This is especially important in places like colonial Virginia. Historic and modern maps are very helpful. A tract of land in a place that was settled early by European immigrants, like the Northern Neck of Virginia, might find itself in three or four different counties before its current location.

<https://markarslan.org/genealogy-public/Head-ColonialVA/Descendants-ColonialVAHead.pdf>

<https://markarslan.org/genealogy-public/DescendantsKey.pdf>

1115. Sarah HEAD (female)

B. [34467]

1116. John Alfred HEAD (male)

B. (4 Oct 1777)|(9 Oct 1787 [27278, 27287, 27292, 34467, 34473])

D. 18 Sep 1847 - Nov 1847 @ (Franklin Co., Kentucky)? [27274, 27299]

M. [1] 20 Sep 1805 @ Franklin Co., Kentucky to [27275, 27274, 34473]

Polly HEAD [#5432]

(daughter of Benjamin HEAD & Mildred LONG)

b. 14 Apr 1787 @ Kentucky? [27299, 34473]

d. after 1850 @ (Franklin Co., Kentucky)?

11161. Benjamin T. HEAD (male)

B. 17 Jun 1806 @ (Scott Co.)?, Kentucky [27287, 27293]


For each event (birth, marriage, and death), the links to the relevant primary source citations will be shown. These references are evidence that either support the conclusion you've reached (date and/or place) or that have something relevant to say about it (even if it is in conflict with your conclusion!). If a person in my tree is mentioned in a significant entry like a will, court minute, deed, etc., I may add that reference to the birth event. Evidence supporting filiation to a parent may be placed here, as well. Note (in this example) that John Alfred Head's descendant ID will appear in the abstract notes file as "[#1116]". His wife Polly Head [#5432] was his 3rd-cousin, a descendant of this same family.)


27276 Court Records: KY, Franklin Co., Will Book 5 (1855-1881)
27277 US Federal Census: 1810 KY, Franklin Co., p. 54
27278 US Federal Census: 1810 KY, Scott Co., p. 640
27279 US Federal Census: 1810 KY, Henry Co., p. 463
27280 US Federal Census: 1810 KY, Henry Co., p. 462
27281 US Federal Census: 1820 KY, Franklin Co., p. 82


27282 US Federal Census: 1820 KY, Franklin Co., p. 86
27283 US Federal Census: 1820 KY, Franklin Co., p. 100
27284 US Federal Census: 1820 KY, Henry Co., p. 182
27285 US Federal Census: 1830 KY, Franklin Co., p. 117
27286 US Federal Census: 1830 KY, Franklin Co., p. 129
27287 US Federal Census: 1830 KY, Franklin Co., p. 133
27288 US Federal Census: 1830 KY, Franklin Co., p. 131

The platform you use for presenting your results (like Trees in Ancestry.com, below) will display the citations. (You may also choose to link to the actual images within Ancestry.)






Edit birth  John Alfred Head

 Edit fact

 Attach media

 **Attach source**

Source citations attached to Birth.

-  US Federal Census: 1810 KY, Scott Co., p. 640
-  US Federal Census: 1830 KY, Franklin Co., p. 133
-  US Federal Census: 1840 KY, Franklin Co., p. 312
-  Court Records: KY, Scott Co., Will Book A (1792-1809), pp. 199-201
-  Deeds & Land Records: KY, Franklin Co., Deed Book H (1819-1820), pp. 451-454

Conclusions

Publishing your genealogical findings with carefully documented primary documentary sources as evidence will greatly enhance the value of your work for future generations. Researching with primary sources (including DNA findings) is much more interesting and rewarding and adds unique value to the stories about your family.