Online Evidence Charts to Help Students Systematically Evaluate Theories and Evidence

Alex O. Holcombe, alex.holcombe@sydney.edu.au

Hal Pashler

Alex O. Holcombe is at the School of Psychology, University of Sydney

Hal Pashler is at the Department of Psychology, University of California San Diego

Refereed paper to appear in *Proceedings of the UniServe Science Annual Conference*, Sydney, 2010

ABSTRACT

To achieve intellectual autonomy, university students should learn how to critically evaluate hypotheses and theories using evidence from the research literature. Typically this occurs in the context of writing an essay, or in planning the introduction and conclusion sections of a laboratory project. A student should distill relevant evidence from the research literature, evaluate evidence quality, and evaluate hypotheses or theories in light of the evidence. To help students achieve these goals, we have created a web-based "evidence-charting" tool (available at *http://www.evidencechart.org*). The main feature of the website is an interactive chart, providing students a structure to list the evidence (from research articles or experiments), list the theories, and enter their evaluation of how the evidence supports or undermines each theory/hypothesis. The chart also elicits from students their reasoning about why the evidence supports or undermines each hypothesis, and invites them to consider how someone with an opposing view might respond. The online chart provides sortable summary views so that one can, for instance, see the evidence indicated to be most important for each hypothesis. Upon completing a chart, the student is well positioned to write their essay or report, and the instructor can quickly provide formative feedback indicating whether the student has successfully reviewed the literature and understands the evidence and theories. These benefits are being evaluated in the context of introductory and advanced psychology classes.

University graduates should be independent thinkers. In today's world, the knowledge needed to succeed in many occupations can change rapidly. Specific content information learned at university frequently becomes outdated or obsolete after a few years (Scardamalia & Bereiter 2003).

Today, a wealth of task-relevant information is often available at one's fingertips through the internet. However, assessing which information is truly relevant to the task question at hand can be difficult. Once the relevant information has been identified, the next step can be even more difficult. The inquiring person should next critically evaluate the information and synthesize it into an overall answer.

Consider an IT manager contemplating which of three types of computers would perform better for a certain purpose. Or, a veterinarian trying to decide which of four possible treatments to administer to a horse with a particular disease. Or a business consultant facing a series of deadlines who wants to know whether drinking coffee or taking naps would be better for his productivity. For each of these questions, there may be no authoritative reference work available that provides the answer. To make an intelligent decision, these professionals must consider what kind of evidence would be relevant to their decision, how they might acquire that evidence, seek it out, organise it, and synthesize it into an overall answer.

These skills of independent inquiry do arise in many university curricula. More precisely, a need for these skills sometimes arises, although the skills themselves are not always taught effectively. The skills are utilized in essay assignments or laboratory research projects. For example, for essays in certain science classes students must examine the research literature to evaluate theories or hypotheses. Laboratory projects also have potential for fostering intellectual autonomy. In a basic laboratory exercise, students are given a set experiment and learning is restricted to understanding a specific experiment and the analysis of its results. However, in cases that foster more intellectual autonomy, students are asked to write an introduction that sets out hypotheses or theories, and in the conclusion evaluate the theories in light of their own results and that of results reported in the literature.

For both a research-based essay assignment and a lab report that engages with the research literature, a student may need to perform the following steps:

1. In response to a question or point of contention, formulate candidate theories or hypotheses
2. Glean relevant evidence from original data or from scientific literature
3. Organise the evidence and evaluate how each speaks to the theories or hypotheses considered
4. Synthesise the evidence and their intepretation of it into an overall answer

In the context of a laboratory report or scientific essay, students should already be performing each of these tasks. It is our experience, however, that students frequently fail to succesfully complete one or more of these tasks. Unfortunately, identifying where the failure occurred can be difficult. Assessments of student work frequently consider only the final product of the process—a finished report or essay. This makes it difficult to determine which steps of the process were done properly and which were not. The difficulty is compounded by the fact that many students do not write clearly. While helping students clarify their writing can sometimes be done with solely the final document, identifying which of the preceding steps went wrong is more problematic. And without focused feedback regarding which steps were not performed properly, many students will persist in their mistakes.

The "evidence-charting" tool described below is designed to achieve two outcomes:

1. Support student performance of the four steps identified above.

2. Create evidence of student performance of these steps, and make it easy for an instructor to assess.

The evidence-charting tool we have created is embodied in a website. The tool is viewable at *http://www.evidencechart.org*, and hereafter this particular software will be referred to as EvidenceChart. It provides a structure with slots in which the student adds information to create an organised summary of their research and some of their thinking. As the student proceeds, the constant presence of the structure reminds the student of what is to be done.

Charting the Evidence

The evidence chart is oriented towards answering an empirical question. It revolves around the candidate hypotheses, relevant evidence, and how each piece of evidence speaks to the hypotheses. The EvidenceChart site has slots for this information in its two-dimensional tabular structure. Each column addresses a particular hypothesis, and each row a particular piece of evidence. Each interior cell of the resulting matrix is the meeting point of a theory with a piece of evidence.

This tabular representation is rather intuitive and has apparently been invented repeatedly over the years. It has been used systematically in communities of intelligence or national security analysts, where it is called the "Analysis of Competing Hypotheses" method (Horn 1999). It has also been used in classroom settings, but reports on its usage are scant. The exception we have found is the Belvedere education project, which includes evidence charts in its Java software for student collaborative inquiry, wherein students created hypotheses, discussed them, and made diagrams as well as an evidence chart to further their inquiry (e.g. Suthers, Toth, & Weiner, 1997). The software does not support online collaboration, but is still available as functioning Java software from the project website. Our effort has been restricted to making a website with easy-to-use evidence charting, plus accessory functionality that assists instructor evaluation and response to what the student has done. By creating a website focused on this relatively narrow enterprise, we hope to keep the programming challenge manageable and maintainable while still having enough functionality for the site to be useful in various contexts.

Our approach is design-based research: implementing and improving our evidence-charting tool in iterative fashion. Following the use of the tool in a university class, we collect feedback from students and instructors and then revise the website and associated instructional material and assessments for the following semester.

In the current iteration, when the student visits *http://www.evidencechart.org*, they begin with an empty evidence chart, as shown in Figure 1.

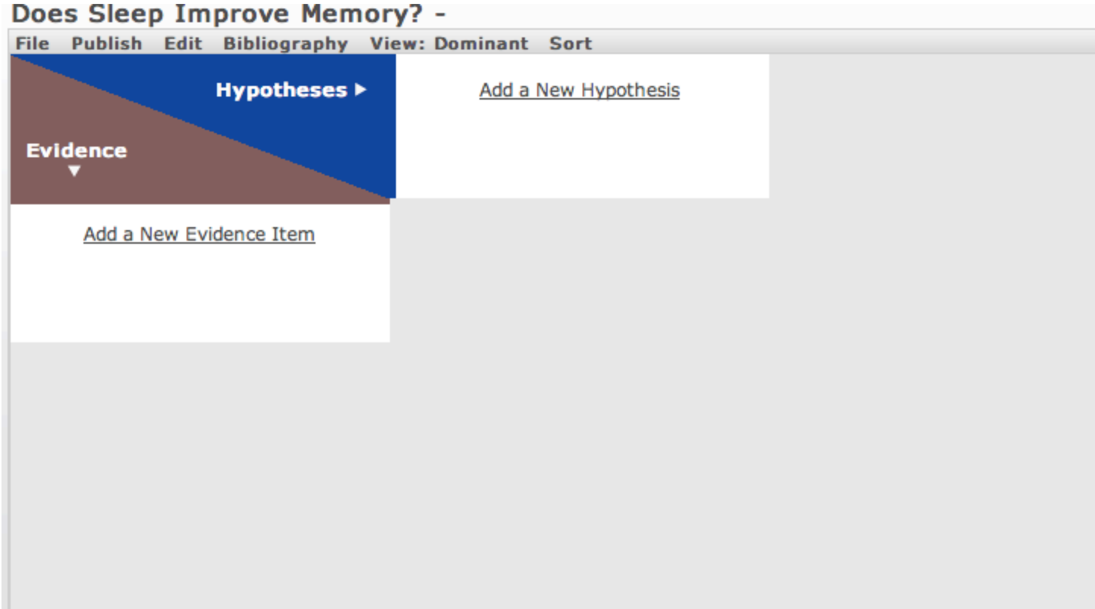

Figure 1. An empty evidence chart immediately after its initial creation and assignment of a title. From www.evidencechart.org

The underlined links shown in the screenshot (Figure 1) indicate to the student that she should add hypotheses and evidence by clicking on the indicated text, after which text input boxes appear and prompt the student to enter corresponding information. As a student does the work outlined in the four steps described in the introduction, she gradually populates the chart. A portion of one such chart is pictured in Figure 2.

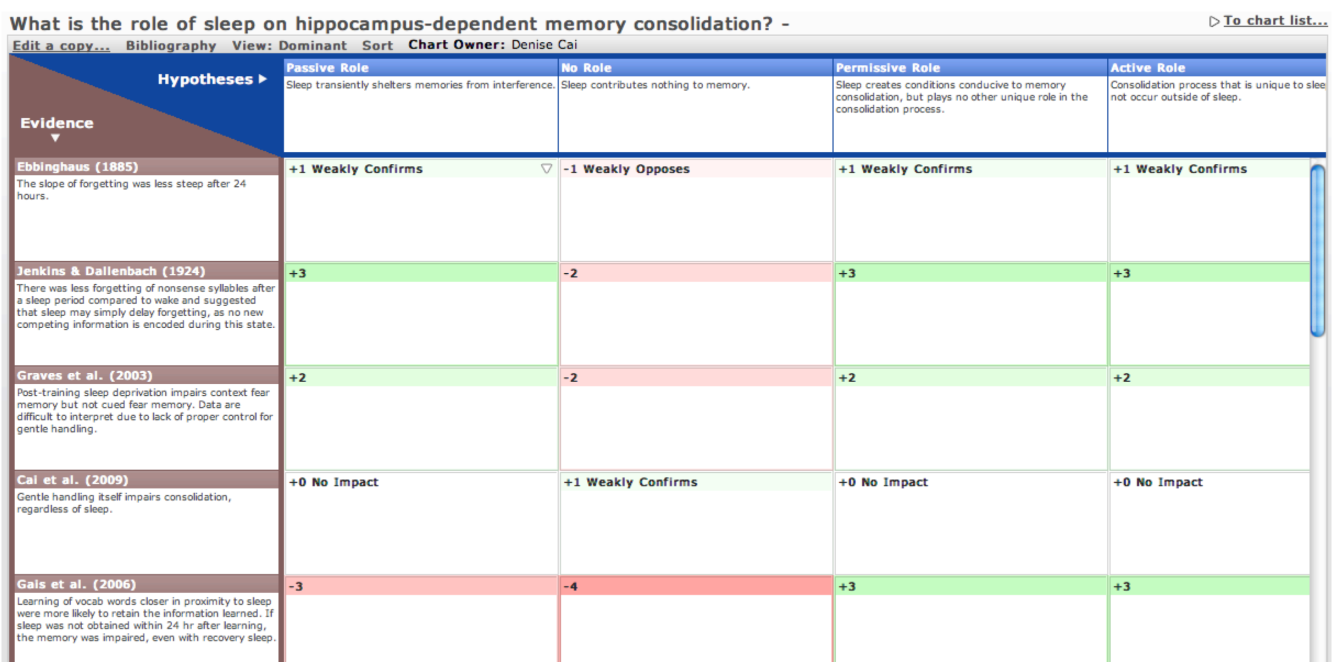

Figure 2. A portion of an evidence chart. The chart was created by Denise J. Cai (UCLA Physiology) and is used with her permission. Evidence (rows, labeled in the leftmost column) and hypotheses (column headers) have been entered, and the degree to which each piece of evidence supports or undermines each theory has been indicated. The student should continue by entering text at each interior cell of the matrix to indicate why the corresponding evidence supports or undermines the corresponding hypothesis. A further aspect is a 'contrarian view' of each cell, in which the student is

encouraged to play devil's advocate and describe the best argument against the position they have taken in this dominant view.

In this chart, each row represents a published scientific article or monograph with results that bear on the question of how sleep affects memory consolidation. Each column describes a different hypothesis regarding the role of sleep in memory consolidation. At the intersection of each row and column, the student should:

- Rate the implication of the evidence for the hypothesis, on a scale spanning "strongly undermines" (color-coded with red) to "strongly supports" (green) the hypothesis.
- Enter a phrase explaining why they believe the evidence supports/undermines the theory. This is termed the "dominant view".
- Engage in 'devil's advocate' thinking by entering a phrase defending the view opposite to what they have indicated in the dominant view. This in entered in an area revealed by clicking on the View menu.

These three functions occur at each cell of the table and systematically coax the student to think critically about the evidence and the hypotheses. The text entered into the contrarian view encourages the student to take another perspective, allowing the student herself to provide the useful and classic 'devil's advocate'. The 'devil's advocate' technique descends from the classic method of Socrates and is commonly used in educational contexts such as law schools; the law professor customarily challenges a student's argument by raising arguments against the student's position. Without some kind of prompting, many students writing an essay or lab report will amass arguments for their position but never think actively about the best arguments against their position. The evidence chart encourages contrarian consideration without the requirement for active intervention by an instructor.

The student's activity described so far is primarily analytic, considering each piece of evidence as an individual. Eventually, the student should shift to synthesizing the evidence and its implications to arrive at a coherent view. Such synthesis of possibly disparate and contradictory pieces of evidence is clearly a subtle enterprise that cannot be reduced to a formula or algorithm. It requires more than simply 'adding up' evidence that seems to be for and against an argument. The evidence chart web application does however provide a small degree of assistance. By clicking in a drop-down menu associated with each column, the student can sort the rows by degree to which he has indicated the evidence supports or undermines the theory. This can be very useful for considering the strongest evidence for or against a hypothesis—particularly for larger charts, such as the full chart excerpted in Figure 2, which contains 20 rows in its full form. A further feature, not yet implemented, may sort the evidence rows by the extent that they discriminate among all the theories.

Using EvidenceChart To Improve Feedback and Assessment

When a student receives a poor grade or mark, the student should be told which aspects of their performance were responsible for the poor outcome. Lab reports and research essays can include several steps before the writing begins and from a poor final product, it can be difficult to judge which steps were at fault. Some students are on the wrong track well before beginning to write, but persist in following their ill-conceived notions or process to a mistaken conclusion. The EvidenceChart webtool makes it easy for an instructor to assess student performance of the suggested steps prior to the writing of a final report or essay. Through the website, user accounts of students in a particular class are grouped together, and class instructors can view the evidence charts they create as part of the class. By requiring each student to prepare an evidence chart, instructors can assess whether a student has found the appropriate related evidence, been able to articulate competing hypotheses, and has some understanding of how the evidence supports or undermines each hypothesis. Thanks to the succinctness of evidence charts, they can do so quickly. Without such a concise format, in large classes it is often impractical to provide individual attention to students prior to final assessment.

A particular advantage of the web-based implementation of evidence charting is that instructors can 'drop in' without the student needing to submit anything formal. In the evidence-charting site, the instructor can add a note for the student indicating which parts appear to be a problem. For a project or essay, rather than have a single deadline corresponding to the final product, students can be required to complete an evidence chart online some weeks before the essay or product is due. As a graded component, the instructor may simply wish to confirm that the student has done something substantial, but for formative assessment can take the opportunity to guide the student with comments on the chart. In addition to correcting students who misunderstand the hypotheses or related evidence, this also curbs the student procrastination problem by making it a requirement that students do some substantial research and thinking well before the final assignment is due.

Current Experience and Prospects

Creating the web application has been a large software development effort, involving many cycles of planning, programming, and assessing the utility and usability of the website. To be truly successful the tool must be very easy and quick to use, or students will resist it. This provides a significant user interface and web programming challenge. As the website has not been stable but rather changed and improved continuously, with intermittent bugs arising in the process, we have not yet mandated that students use it in any class. However, for two semesters the site has been presented to the students as a tool that could benefit them and which they may use if they wish. We have also used it in our own unrelated scientific research to evaluate the viability of various scientific hypotheses.

In using the tool ourselves for professional scientific research, we have been surprised by its effectiveness at eliciting new critical insights. For example, one of us studied a particular visual illusion for two years and formed various opinions of the theories that have been proposed to explain the illusion. Simply to test out the website functionality and ease of use, it was decided that an evidence chart regarding the topic would be constructed. The process prompted focused consideration of how each piece of evidence could or could not be reconciled with each theory. This proved very productive, as several novel insights were gained. Although previously much of the evidence had been considered extensively in light of one or two theories, never had each piece of evidence been considered for each. We believe that most students as well as working scientists also do not usually approach a problem very systematically. Many scientists know that there is nothing like writing an article or grant to force onself to consider a theory more carefully. However, writing prose can be daunting and considerations of exposition, clarity, and organization can become prominent before one gets through very much evidence. In contrast, the very limited space provided in the cells of an evidence chart elicits a short phrase or two accompanied by careful thinking. The blank space of those entries where the evidence has not been fully evaluated are persistent reminders that one has been negligent. The existence of such omissions are easily forgotten or never even realized without an evidence chart. Furthermore, the result of the process provides a product that facilitates synthesis of the evidence. In traditional prose format, synthesis seems more difficult. One reason is undoubtedly the limitations of working memory: it is simply hard to keep in mind the points made in many different paragraphs regarding how a half dozen pieces of evidence relate to three different theories.

Student feedback on the usefulness of the tool has been limited to date, but encouraging. At the University of Sydney, the tool has been presented to students in a large introductory psychology class consisting mostly of first-year students, to fourth-year (honours year) students working on a year-long research project, and to a few postgraduate students in psychology. Evidence charting was described as entirely optional and it seems that only a small proportion chose to attempt this additional activity. Feedback has been solicited via prominent hyperlinks on the website and electronic surveys emailed to many of the students. Responses have been few, limited to a dozen or two, and have consisted of two types. First are reports of problems or perceived problems with the functionality of the website. For each of these negative reports, together with our programmer we have been able to quickly resolve the issue. All other comments have been positive and have often been provided by a person who also complained about a possible bug with the site. When receiving a complaint, we take that opportunity to engage the person and ask about the site's general utility. Here are a few of the comments we have received:

Because I have reasonably slow internet, occasionally the program had trouble saving the information I had just added. Which was mildly annoying, but overall it was a really awesome tool. I'll definitely use it again when I restart my degree in a few years. =)

I created an account and successfully started using EvidenceChart - it is seriously amazingly helpful because Microsoft Word and Excel are absolutely crap for this sort of thing..... And like I said, this is amazingly helpful in sorting out the literature! Thanks for getting this out to us :)

A PhD student who we commissioned to test the site by making a chart associated with her doctoral work provided the following feedback:

It makes me think of the contrarian view, which is great! While I think about this all the time, it's actually really helpful to verbally articulate it and then document it! It's also been helpful in dissociating between the strength of confirmation/opposition for a theory vs rigorous/"well-doneness" of a study, as mentioned before. I'm sure it'll help me gain more "ah-ha" moments as I start working on a less familiar topic.

This doctoral student, together with others, mentioned the difficulty of choosing the best level of granularity for the evidence. In the case of preparing an evidence chart for a scientific essay comparing theories, for example, should each row refer to an individual experiment, an entire scientifc article, or a set of scientific articles containing similar experiments? This can be difficult to know before most of an evidence chart has been constructed. When the appropriate level of granularity has been chosen, certain pieces of evidence may be highly related; for example, they may all bear on a single aspect of a hypothesis. Ideally, this evidence should be grouped together or perhaps be part of a larger hierarchy. However, it has been difficult to envisage software support for this without making the user interface substantially more complicated. Our aim is to keep to a simple design that a novice can use immediately after nothing more than a one or two-minute explanation.

As the software has been tested by many dozens of student volunteers and crashes and bugs are now rarely if ever encountered, we are ready to move to the next phase of the project: mandating that students create an evidence chart prior to writing their essay or lab report, and providing them with rapid formative feedback on the basis of the chart. Following this, there are plans to modify the software to allow for collaborative group editing of evidence charts. This will allow groups of students to work together on the chart (using their individual logins), allowing them to learn from each other, even across large distances, and more independently from the instructor.

Full-formed prose writing is clearly not an optimal format to start with when planning a critical essay. It is not suprising, then, that long before evidence charts and computers were invented, there were other techniques that students used to plan their essays. For example, many scholars and students, especially in the humanities, put bits of information on individual small cards or "index cards". Typically, one piece of evidence is written on each card, similar to the individual rows of an evidence chart. After the evidence is amassed, the cards are assembled into a linear or two-dimensional array that has some sort of correspondence with the argument or composition being planned. The potential to create practically any structure with this technique means it is suited to any purpose. At the same time, however, it does not provide a guiding structure for a student who is not yet a master of the process. Similarly, concept-mapping and mind-mapping are very flexible but provide few relevant structural constraints. Argument maps are highly structured and very promising for concisely representing arguments but require extensive training to learn (vanGelder 2002). An intermediate between these extremes, something like evidence charts, may eventually take hold as a helpful tool for students and professionals. The added interactivity and limitless functionality possible in internet-connected software will undoubtedly be an intimate part. The evidence-charting tool is useful now and we hope it is moving in the right direction to help students and scholars work efficiently, systematically, and think critically.

References

Horn, R. (1999). Analysis of Competing Hypotheses. In *Psychology of Intelligence Analysis*. Center for the Study of Intelligence, CIA.

Scardamalia, M. and Bereiter, C. (2003) Knowledge Building. In *Encyclopedia of Education*, MacMillan.

Suthers, D.D., Toth, E., & Weiner, A (1997). An Integrated Approach to Implementing Collaborative Inquiry in the Classroom. In *Computer Supported Collaborative Learning '97*, December 1997, Toronto.

van Gelder, T. J. (2002). Enhancing Deliberation Through Computer-Supported Argument Visualization. In P. Kirschner & S. Buckingham Shum & C. Carr (Eds.), *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*. London: Springer-Verlag, pp. 97-115.

Acknowledgments

This work was supported by the US National Science Foundation (Grant BCS-0720375 to H. Pashler, and Grant SBE-582 0542013 to the UCSD Temporal Dynamics of Learning Center) and by a collaborative activity grant from the James S. McDonnell Foundation.