Purdue University

Purdue e-Pubs

Charleston Library Conference

Playing the Odds: Pascal's Wager and Decision Making in the Long Scholarly Conversation

Douglas Black
Northern Michigan University, doblack@nmu.edu

Playing the Odds: Pascal's Wager and Decision Making in the Long Scholarly Conversation

Douglas M. Black, Collection Development Librarian, Northern Michigan University

Abstract

This talk explores some of the philosophical issues surrounding the nature of the scholarly record and current challenges in academic libraries. The discussion arises from a 2011 *Chronicle of Higher Education* feature on a widely influential 1979 article entitled "Prospect Theory: An Analysis of Decision Making Under Risk," which proposed a new model for how people assess risk and weigh decision factors. The theory has been applied to dozens of disciplines and cited thousands of times, and it has applications in academic collection development as well as everywhere else. It addressed the limitations of Utility Theory, which grew out of Pascal's Wager (i.e., it's safer to bet on the existence of God), but didn't adequately explain how people—gamblers and insurance buyers, for instance—actually weigh risks and make decisions. The pace of change and new demands facing libraries offer the opportunity to ask new questions about the nature of the scholarly record. Increasing ubiquity and transience of information, along with rapidly shifting notions of authorship and ownership, offer some interesting questions about the role of academic library collections in scholarly communication as a whole.

Introduction

In the academic library, we all know how difficult it's become to keep up with increasingly rapid developments in information, and our current environment challenges our basic notions of what constitutes the human record (disclaimer: terminology here deliberately ignores the difference between "human record" and "scholarly record"). We feel increasingly like we're doing a lot of betting, and I'm going to explore here the nature of that betting in the context of what we've traditionally thought of ourselves as doing. One way of understanding the problem is to consider ubiquity and transience.

Ubiquity

Here's what we usually think of when we think of the scholarly conversation over the centuries. I've often used it in instruction sessions to help students think about attributing and synthesizing source material. It's the standard view of how our understanding of the solar system developed.

1. Ptolemy, the 2nd century Alexandria astronomer, formulated a geocentric theory of the universe, in which the sun and the other planets orbit around the Earth.

2. Copernicus, in the 16th century, came up with the heliocentric theory that all the planets orbit around the Sun.

3. In the early 17th century, Galileo (1564–1642) aligned his views with Copernicus but couldn't fully prove it by observation.

4. Around the same time, Johannes Kepler (1571–1630) worked out some laws of planetary motion, continuing the work of his mentor Tycho Brahe.

5. Late in the 17th century, Isaac Newton did his work on gravity and showed its consistency with Kepler's work on planetary motion, confirming Copernicus and Galileo.

This sequence is one example of the long scholarly conversation into which we invite our students. We invite them into scholarly discourse and begin training them to think about how ideas connect, how we know what we know both as individuals and as humans, and how we advance what we know both individually and collectively (Berrett, 2012). As we draw students in and forward, we engage them in creating new knowledge. That process, fundamental to the academy, lies at the heart of the problems facing the academic library, so long considered the repository of the scholarly record.

In 2011, the *Chronicle of Higher Education* ("The reach of 'Prospect theory'") charted the spread of an article called "Prospect theory: An analysis of decision under risk" (Kahneman & Tversky, 1979). The *Chronicle* feature showed the increasing rate of citations in a number of diverse disciplines. At the time, the *Chronicle* estimated that the Kahneman and Tversky article had been cited some 8,000 times.

However, we can't collect all of that information, especially not at below R1 level—and R1s can't keep up either. In the small regional university library, where we explicitly focus on providing a representative, not comprehensive, collection, we can't keep up with what we feel we ought to be collecting. However, consider the academic trend (in the humanities, mainly) in the last few years toward emphasizing primary-source documents. Consider the proportion of the human production of knowledge that we've recorded and preserved over time, against the proportion of materials that we've never captured for preservation, especially primary materials: everyday shopping lists, letters, handbills, theatre programs, petitions, windshield flyers, all the gray literature that's gotten away, receipts, S&H Green Stamp booklets, chain letters, all the e-mails...when we consider everything we've created against what we've managed to preserve, what becomes clear is that we've only ever preserved a fraction of it to begin with (Anderson, 2011).

When ideas spread comparatively slowly and publication technology was comparatively expensive, it seemed realistic to think that we could keep track of human knowledge. But now, a given scholar in a given field can't possibly read and absorb everything written in her field anymore, and studies are showing that scholars are reading and citing differently now (McClanahan et al., 2010). Put simply, it's too much. When information products were valued in part based on their scarcity, then especially, collecting on behalf of a community made sense; first, because the collection itself represented the record of the community; second, few if any community members could afford to gain access to all of it on their own; and third, because even though it was growing, the sum total of the

materials being produced appeared that it would always be a relatively finite, comprehensible set of objects.

So here, let's take another look at astronomic research and advances over the centuries. The most important clue to the truth lies in the millennium and a half between Ptolemy and Copernicus, which are just the most commonly recognized names. There were more, many more:

- Ptolemy 100–178 AD
- Abu Mashar, 805 or 806–886 AD
- Donnolo, Shabbetai 913–ca. 982
- Alhazen, 965–1039
- Petrus Alfonsi 1062–1110?
- Roger of Hereford fl. 12th cent.
- Gervase of Canterbury ca. 1141–ca. 1210
- Grosseteste, Robert 1175?–1253
- Levi ben Gershom 1288–1344
- Ibn al-Shāṭir, ʿAlī ibn Ibrāhīm, 1304–1375
- Heybech, Nikolaus von fl. 1389–1444
- Al-Ahdab, Ishak ben Shlomo fl. 1396–ca. 1429
- Zidek, Pavel 1413–1471
- Moncada, Guglielmo Raimondo de ca. 1450–ca. 1489
- Copernicus 1473–1543
- Martin Luther 1483–1546
- Leoninus, Albertus 1534–1614
- Brahe 1546–1601
- Galileo 1564–1642
- Kepler 1571–1630
- Newton 1642–1727

Obviously, we know even these names only because some form of their work, or reference to it, has survived. But not many people outside the history of science, astronomy, or other related fields would recognize all of them.

Add to this ubiquity the much more recent decentralization of information creation, with users becoming creators as well; now that everyone is a producer of information, our sense of finity and control in the bibliographic world is truly obliterated. But perhaps that sense of finity and control was illusory all along. How many more people were active in astronomy over the centuries whose work has not survived or has not been widely known enough to be catalogued? How would we know?

So we probably never did collect enough to preserve the human record in anything like the comprehensiveness we'd like to think. That realization is emerging only now that we're aware of how much we can't collect because of our current circumstances and how rapidly they have developed.

Transience

Following indirectly from information ubiquity is its transience or impermanence. For example, one recent work about YouTube constitutes a "book" entirely online (Juhasz, 2011), under the ongoing control of the contributors. It's not only a treatment of YouTube, but a manifestation of its own subject. There's an interesting wrinkle here that's not immediately evident. This work's significance is less about its form or its YouTube focus than the questions it forces regarding peer review and evaluating scholarship: The content of this online book isn't fixed. The students own the pieces they've produced and can remove their material any time. Juhasz knows that's likely to happen and notes that replacement content can always be added to fill the gaps (Parry, 2011).

Our most traditional methods of assessment, on the other hand, have always assumed static content, in that a printed volume read 2 years ago will still have the same text and illustrations next year. We often review new scholarly work by placing it in context with the existing literature. As material and content become more ubiquitous and distributed, however, distinctions blur; If the work can change so substantially, beyond the usual range of updating editions, what are we really evaluating? The enduring value of the content, the impact of its production as a

phenomenon, or the contemporaneous quality of its conception and execution? If "the book is so temporary" (Parry), then on what do we rely for assessing the quality of the scholarship? Five or ten years after its initial review, its composition will probably have changed, so what might its peer-reviewed status mean over time? Is it simply a new edition that calls for a new peer review? If it's intended in the first place to be impermanent, should review take into account the work's purpose? Would the fact that it has changed later on mean that it had successfully fulfilled its purpose, even if the content were no longer the same? It is more akin to a juried performance, or is it something else altogether? If we consider it critical material, then later on, when it becomes part of the existing literature, how does its transient nature affect our evaluation of still newer work to be compared to it? How do we begin reenvisioning our methods of evaluating scholarly work created under new publication models like this one?

When the primary material is ephemeral, not merely uncertain of access due to format, but ephemeral in its content by explicit design, what happens to our own critical distance? What are the broader implications for scholarly communication over time? What distinctions do we draw between object and form or method? What remains, and in what ways does it matter what remains?

For scholars of art and architecture, the latter questions aren't new. For scholarly communication in general, Juhasz (2011) is fascinating partly because it raises so sharply the question of what a work is and demonstrates how fluid even basic precepts can become. Those precepts might have always been rather more fluid than we've been aware, simply because of the nature of publishing and the pace of advancement until recently.

Which brings us back to "Prospect Theory" (Kahneman & Tversky, 1979). This article is important because it presented a brand-new theory on how people calculate the likelihood against the consequences of various alternative outcomes. Utility theory arose out of Pascal's Wager. In a possibly outrageous

oversimplification, the question Pascal posed was whether to believe in God. The reasoning goes like this: If I choose to wager against God existing, I can have a good time on Earth but risk eternity in Hell if I'm wrong. If I'm right that God doesn't exist, I lose nothing in the long run. If I choose to wager that God exists, I stand to gain eternity in Heaven in return for living according to God's law. If I'm wrong that God exists, I don't gain anything in the long run, and I've given up a lot of pleasure on Earth. Rationality dictates that I wager for God's existence, since the utility of the Heaven prospect after a righteous lifetime outweighs the utility of pleasure on Earth with the possibility of eternal Hell.

Utility theory, however, didn't seem to account consistently enough for how people actually made decisions that carried risk (Kahneman & Tversky, 1979, p. 263), such as in gambling and purchasing insurance. When probability and rates of return are involved, such calculations as Pascal's aren't so simple. "Prospect Theory" closely examined probability and perceived probability via a number of experiments and games investigating how people responded to specific scenarios.

One example from Kahneman & Tversky (1979) is the concept of probabilistic insurance. In the standard system, you purchase home insurance against the risk of fire. The insurance company is betting that nothing will happen to the home, and you're betting that something will happen. If the premium is as much or more than the cost of a potential loss, purchasing the insurance doesn't make sense. Put this way, utility theory seems adequate to explain the decision.

The idea of probabilistic insurance illustrates more clearly how people actually think about insurance. The homeowner's choices are modified to introduce a second layer of risk independent of the risk of fire: Pay half the premium, and there's a 50% chance that a loss will be covered. If it is covered, then you pay the other half of the premium and the company pays for the loss. If it isn't covered, you get your money back and pay for the loss yourself. What determines whether the loss is covered is unrelated to the risk of fire; the article's example is that if the fire occurs on an odd-numbered day of the month, it's covered, so

you pay the other half of the premium and you're covered. If the fire occurs on an even-numbered day, you get the premium back but no coverage (pp. 269–270). Kahneman & Tversky explain that "expected utility theory...implies that probabilistic insurance is superior to regular insurance." However, regular insurance eliminates all the risk of loss from the fire, whereas probabilistic insurance actually introduces additional risk (p. 270).

In short, we might believe we're making a choice between clear alternatives, with careful consideration of utility and likelihood, but very often, our choices merely minimize risk instead of eliminating it—and sometimes even increase it. We often do think in terms of gambling against the unforeseeable, but we don't always realize that that's what we're doing. The way we make decisions, in other words, isn't always what we think it is.

In the academic library, our traditional method of selecting and acquiring was in case something would be needed, based on our particular institution's mission and our knowledge of historical and/or current instruction and research. The large research libraries were thought of as the repository of the human record, preserving collected human knowledge for posterity. In smaller institutions, we've aimed for representative, not comprehensive, collections, partly because comprehensive collecting has been completely beyond our fiscal reach. The idea was to purchase what we could afford and what we really needed, creating an adequate basis for ongoing scholarly activity on our own campuses. Yet we still thought of our libraries as the official repository of human knowledge, and in acting conservatively just in case, we all too frequently engaged in what we thought was Pascal's Wager.

For example, cancelling newspaper microfilm because the title was available via an aggregator and having the publisher cancel its contract with that vendor years later. We might be said to have made what we thought was a simple utility decision, choosing the greatest benefit from amongst the alternatives by saving money and providing better access. The decision turns out later to have been a gamble, and a losing one at

that, given the astronomical cost of acquiring the material retrospectively. But this particular losing gamble is less a matter of underestimating risk than of not understanding that it existed in the first place, which is exactly what happens when we purchase insurance thinking that we're eliminating all risk. The decision appears to be a simple utility-theory calculation. However, Pascal's Wager identified a positive utility for only one decision/outcome set, with the others all neutral or negative. Deciding where to invest resources in a zero-sum situation, on the other hand, offers both positive and negative utility for each decision/outcome pair. The various apparent risks and benefits don't always weigh clearly one way or another on the scales.

What we haven't understood in the academic library is where we inadvertently introduce risk in our decision making. The changes we're now seeing in the publishing industry are making some of those risks clearer, and most of us grasp them better now; we know more about what questions to ask when allocating resources or negotiating for access to electronic materials in particular. We know now that access via an aggregator is less reliable than direct subscription, even if the aggregator appears more cost-effective. We know now that as print journals cease or move to online only by the dozen, we're getting nervous about retaining access over the long term. We know now that there's even a lot of uncertainty underlying our purchases of materials even with explicit perpetual-access rights directly from the publisher, and we know that those purchases are still something of a gamble. We don't have any way of hedging our bets, but at least we now know that we're betting in more complex ways than we thought.

Understanding how we actually make decisions in uncertain cases is what Kahneman and Tversky were getting at. Reading that article first probably wouldn't have changed our *New York Times* decision, since nobody knew what the risks might be, but reading it now helps us understand the decision. It also helps us examine new decisions with a better sense of how we're likely to weigh the factors and what unknowns to think about going forward.

In the world of academic collection management, we have decisions to make every day in times and conditions of what we consider grave risk in the broader cultural arena, and in the context not only of what we do historically, but of our historical sense of what we do. In the academic library, we all recognize that we can't adequately absorb and respond to current and future changes either physically or fiscally. We can step out and lead in some ways, but on our campus, at least, our primary mission is to support our own curriculum and research, which means we have to retain some significant measure of reactivity.

We all hear the constant calls to embrace change better in order to preserve the profession, the institution, and its relevance, all alongside a sense of impending disaster for the human record. I wonder whether we might try asking different questions. We can no longer afford to preserve the human record adequately, but it's not necessarily clear that we ever actually did that. And given the whiplash changes in information formats, production, usage, and lifecycles, what happens if we try to conceive instead of our ongoing role in scholarly communication as that of mediators and facilitators, rather than as keepers or curators? Selecting materials has always meant mediating, and providing access has always meant facilitating; that's part of what we know we've been doing all along. We've also long felt that our central role was to preserve the human record. What I suggest here is that the nature of the human record may be fundamentally changing, and it may have been changing for as long as we've had the concept of a human record. We're noticing it only now that change has accelerated and accelerated so rapidly. Adapting to such changes means accepting some changes in our identity and self-image, but maybe only the parts that have been illusory or ephemeral all along. Choices remain, and to make them as effectively as possible, we should consider that our decision making involves more uncertainty than we've long thought. With new understanding of how we make our choices, we might need to accept that a certain amount of letting go is the best way to hold on.

References

Anderson, R. (2011). *Why we do what we do, and what we should be doing instead*. Address presented at Michigan Academic Libraries, Grand Rapids, MI, May 6.

Berrett, D. (2012, October 8). Habits of mind. *Chronicle of Higher Education*. Retrieved from http://chronicle.com/article/Habits-of-Mind-Lessons-for/134868/

Juhasz, A. (2011). *Learning from YouTube*. Cambridge, MA: MIT Press. Retrieved from http://vectors.usc.edu/projects/learningfromyoutube/

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–292.

McClanahan, K., Wu, L., Tenopir, C., & King, D. (2010). Embracing change: Perceptions of e-journals by faculty members. *Learned Publishing 23*(3), 209–223. http://dx.doi.org/10.1087/20100304

Parry, M. (2011, February 20). Free "video book" from MIT challenges limits of scholarship. *Chronicle of Higher Education*. Retrieved from http://chronicle.com/article/Free-Video-Book-From/126427/

Streitfeld, D. (2012, February 22). Amazon pulls thousands of e-books in dispute. *New York Times*. Retrieved from http://bits.blogs.nytimes.com/2012/02/22/amazon-pulls-thousands-of-e-books-in-dispute/

The reach of "Prospect Theory." (2011). *The Chronicle of Higher Education*. Retrieved from http://chronicle.com/article/The-Reach-of-Prospect-Theory/129687/