Signal and Noise: The Power of n-Dimensional Query and Education

Jonathan Cornwell
The Center for Internet Research
jrc@tcfir.org

W. Reid Cornwell
The Center for Internet Research
wrc@tcfir.org

Abstract: This paper argues that fundamental problems in the way that knowledge is represented and accessed on the Internet prevent the Internet from achieving its full, positive potential in education. Among these fundamental problems are intrinsic deficiencies in keyword search technologies, deficiencies that place undue burdens upon student, teacher, and researcher alike. Vannevar Bush's 1945 article "As We May Think" is used as a framework for discussing in a historical perspective how knowledge representation, the growth in human knowledge, and the means by which we find and relate knowledge continue to confound our use of knowledge. Correlation-based n-dimensional search technology is offered as an important advance over keyword searches.

Introduction

The Internet presents us with an essential conundrum, a "trichotomy " of sorts: an *ideal* as the greatest library of human knowledge yet created; a *reality* as a vast, complex, and raw conduit and milieu of human expression both positive and negative; and as a still-untapped *potential* to revolutionize important human activities including education. Against this evolving background are the continuing struggles of the education system to cultivate the kinds and degree of knowledge and skills that society requires, and an array of imperatives ranging from the mandates of No Child Left Behind to the need to redefine America's place in the global economic, political, and cultural landscape driven by a seemingly ever-accelerating pace of change. It is an irony of our time – an irony lost on few people – that the Internet is both a source of these problems and a key to many of the solutions we might envision.

Clearly, the role of the Internet – herein used as a cover term for all the various information and communications technologies we have today and their contents – in education is still very much up for debate. However, as this conference demonstrates, there is also an understanding of a kind of inevitability in the Internet playing a role in education if for no other reason that the "demand side" of the system: learners seeking to satisfy their own needs from what the Internet has to offer. Matching the ideal, learners are using the vast resources of the Internet to discover and incorporate knowledge unavailable to even the most privileged students a generation ago, and inventing new means of knowledge sharing and expression. In the grittier reality, learners struggle with the burden of separating the "wheat" of authoritative knowledge from the "chaff" of misinformation, of discovering those few units of knowledge cogent to the task in an expanding universe of knowledge, and, in the worst case, resorting to plagiarism and other forms of cheating in the easy markets of the Internet. While many of the fundamental technologies of the Internet arose out of research and development at education and government institutions, the juggernaut of the Internet was born in hands and minds of individuals who sought to do *something* with it, profound or trivial.

The agents of institutional education – teachers, researchers, schools, government – have not necessarily been willing adopters of the Internet although there is now almost a full generation of teachers that could scarcely imagine not being able to download a worksheet, a lesson plan, a current event article, or blog with other teachers on the problems of classroom management. How many schools do not have a website or a computer lab in the U.S.? How many schools lack any sort of connection to the Internet? These are rhetorical questions because the new norm is connection, representation, and participation on the Internet. But these current adaptations are insignificant in

comparison to way in which the Internet is woven into the lives of the "Net Generation" (Tapscott, 1998; e.g. Barnes et al., 2007) and trivial relative to the potential. Regardless of what the institutions of education may do relative to the Internet, the Internet will play a role in education because learners will make it so as they have done so in the past and do at this very moment. The premise here is that we are beyond the debate about whether or not the Internet may play a role in education; it does and it will.

A Problem Greater than Education

For institutions and learners alike, a question that still lies before us is how we may harness the ideal, tame the reality, and realize the potential of the "network of networks" for all human activities but especially education. The possible answers to this one question are innumerable and, in some ways, even unimaginable, as unimaginable as the World Wide Web was to Vannevar Bush in his article "As We May Think" (1945). But many of the problems which confound us in our quest for the potential of the Internet are old and were clearly understood by Bush (1945):

> "There is a growing mountain of research. But there is increased evidence that we are being bogged down today as specialization extends. The investigator is staggered by the findings and conclusions of thousands of other workers—conclusions which he cannot find time to grasp, much less to remember, as they appear. Yet specialization becomes increasingly necessary for progress, and the effort to bridge between disciplines is correspondingly superficial.
>
> The difficulty seems to be, not so much that we publish unduly in view of the extent and variety of present day interests, but rather that publication has been extended far beyond our present ability to make real use of the record. The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships."

For this quote to have its fullest impact, we must make a tangential argument that the tasks and processes of scholar or scientist engaged in research are substantially the same as those of a learner. That is, discovery and exploration are learning processes. The similarity is especially strong when the researcher reviews the literature of a research frontier as it constitutes the acquisition and understanding of the known – that is, knowledge which already exists in the record – that is personally unknown to the researcher. We must also understand that what constitutes a knowledge "frontier" is in reality at least three kinds of things: that which is unknown to humanity, that which is unknown to the individual, and that which is misapprehended that, under different conditions of observation and thought, have new meaning and import. The learner-as-student and learner-as-researcher play similar roles and, more importantly, confront similar challenges in relation to the record of human knowledge, including 1) navigating and utilizing the record of knowledge, 2) expanding the frontiers of his or her personal known, 3) understanding and evaluating new knowledge, and 4) placing new knowledge into proper relationships in order to continue the process. By drawing this parallel, our goal is to argue that achieving the potential of the Internet in education depends, in part, upon the nature of human knowledge, a problem that confounds more than just the student but also the researcher and anyone else who must make use of knowledge greater than that which they already possess.

In the quote above, Bush highlighted the size and rate of the growth of knowledge, the increasing specialization of "workers" and the knowledge they produce, and the methods by which knowledge is published to the record as the sources of problems which resulted in "superficial" connections between disciplines and disciplinary knowledge, and the "(in)ability to make real use of the record". Bush characterized the means for using record of human knowledge as antiquated. While a great deal has changed in information technologies since 1945, most of the paradigms extant in 1945 for organizing and accessing information predominate today. Bush continued (1945):

> "…for we can enormously extend the record; yet even in its present bulk we can hardly consult it. This is a much larger matter than merely the extraction of data for the purposes of scientific research; it involves the entire process by which man profits by his inheritance of acquired knowledge. The prime action of use is selection, and here we are halting indeed.
>
> Selection, in this broad sense, is a stone adze in the hands of a cabinetmaker.

> The real heart of the matter of selection, however, goes deeper than a lag in the adoption of mechanisms by libraries, or a lack of development of devices for their use. Our ineptitude in getting at the record is largely caused by the artificiality of systems of indexing. When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass. It can be in only one place, unless duplicates are used; one has to have rules as to which path will locate it, and the rules are cumbersome. Having found one item, moreover, one has to emerge from the system and re-enter on a new path.
>
> The human mind does not work that way. It operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain."

Here is where Bush, and we, gets to the meat of the argument. Despite innumerable advances, those means which presently exist for organizing and using the record of human knowledge are still largely "artificial" to the traits and tendencies of the human mind. Alphanumeric indices (including keyword searches) coupled with often highly-specialized ontological/taxonomic structures are still the order of the day, and these structures inhibit as often as they promote recovery/discovery and understanding. In addition, the Internet as a whole is "noisy". That is, the Internet lacks an over-arching structure for recovery/discovery (notwithstanding infrastructural elements such as Domain Name System, and keyword-based Internet search applications such as Google), and is thoroughly heterogeneous in the nature of its content. The Internet may also be seen as "paracrystaline", exhibiting pockets of high-order but short-range organization. Associative structures do exist at varying scales and scopes, e.g. hyperlinks, but suffer from a variety of limitations, including the growing prevalence of authoritative knowledge being published in PDF form, and author perception of relatedness. And, of course, the quantity of knowledge on the Internet, and the proportion of the whole of human knowledge on the Internet, continues to expand rapidly (an estimated 5 exobytes of new data in 2002 (Lyman & Varian, 2003) verses 1.5 exobytes of new data (Lyman & Varian, 2000)). Taken together, the qualities of knowledge as represented in the Internet represent a high-noise, low-signal environment more akin to the structures of landfill than a library; "noisy" relative to the highly contextual/situational intentionality of an individual at any given moment in time. So, while the Internet and its underlying technologies represent a great advance forward in terms of ubiquity and scale of access, the resulting "pile" into which human knowledge has been placed is, at best, difficult to use effectively and efficiently to the degree with which we require. However, we submit that this relative disorder is symptomatic not of the Internet per se, its traits and organizational schemes, but of the state of knowledge itself. The advent of the Internet has merely made more obvious preexisting flaws in the qualities of our knowledge, flaws which were obscured when our principle interface with knowledge collections were within highly-rational and ordered environments like libraries and archives.

In a parallel argument, education at all levels involves the (re)transmission of knowledge. More specifically, at the level of kinds of education we provide to children in our culture, the aim is the cultivation of those abilities and knowledge which we believe are necessary to produce well-enculturated and well-adapted adults. What contents must be (re)transmitted, as well as the methods and conditions under which the processes of teaching and learning might operate best are far from settled, and may never be as the socio-cultural conditions against which we determine what "well-enculturated and well-adapted" mean change over time. Beyond the threshold of those basic skills and knowledge which enable higher-order learning, the state and qualities of our knowledge representations and structures become increasingly important to learner and teacher alike. These issues hardly come into play for the learner in kindergarten where basic literacy and numeracy are the focus. However, for the kindergarten teacher and the institutions of education seeking the best tools and methods for teaching basic literacy and numeracy to kindergarten students, the nature of knowledge and the Internet does have a clear and present impact. And for the high school or college student, who may be tasked with high-order research and critical thinking, the nature of our knowledge and the means by which it may be accessed is of manifest importance. Therefore, it is important to realize that the problems, meaning, values, and valence of these issues are highly contextual from nearly every starting point one might engage the nature of our knowledge and education.

Furthermore, our biological capacity to know and understand is relatively fixed, and further proscribed by time, task, and bio-pscyho-social variation. With little reason to believe that the pace of change and the growth in our knowledge will abate, our reliance on knowledge resources like the Internet can only increase. Skills and the knowledge to use knowledge are necessary but not sufficient; innovation in the nature of our knowledge, the way it is represented and structured, as well as the affordances for interacting with it, is necessary as well. For all of these reasons, and more we have not discussed, we argue that we are in desperate need of new information technologies,

and we assert that problems and solutions which appear either to pertain to education or to domains outside of education have a much greater and inclusive scope.

As a final note on this section, rather than a historical curiosity, the observations of Bush offer an important perspective by reminding us that what may seem like new problems are not, and that there is a degree of illusion to the progress we assume we have achieved in the management of our knowledge. In discarding these illusions of progress, we may in fact discover solutions at scales and scopes so fundamental as to be easily overlooked. In this vein, we should consider a statement which borders on the obvious: a search (or "selection" in the words of Bush), whether conducted in a library, a corporate database, or the Internet, is an intentional act predicated on the need to learn *something*. As such, a search represents an important nexus within the problem of education, both in the context of institutional learning and in such activities as "lifelong learning" and "critical thinking" which, in the past decade, has taken on even greater importance as a goal and product of institutional learning. With this in mind, we will now move to the focus of this paper with a discussion of the keyword search and a new candidate technology for discovery/recovery which we believe represents an important advance for the education domain and beyond: correlation-based n-dimensional queries.

Noise: The Problem of the Keyword Search

Consider the keyword search, a technology that well-precedes the advent of computers in all of the paper-based indices created to catalog information. All of the dimensionality of knowledge is flattened into the single-dimension relationship of membership to a common set based on words out of context and presented in a (linear) list arranged by "relevance". The richness of such searches are proscribed by the knowledge base being consulted, whether the knowledge base is the publically available contents of the Internet or the more limited domain of a fee-for-access scholarly database. More rarely, one might be presented with a search tool that consults multiple knowledge bases simultaneously. And even more rarely still, one might have at his or her disposal complex knowledge mapping and search software that offers more than keyword-based set membership. While more complex tools exist, no other method of recovery/discovery is more ubiquitous than the keyword search. Thus, despite yeoman's work across decades, we are still using "a stone adze" for "selection".

Keyword searches depend on the individual's knowledge of the jargon of the subject to be explored. Let us be clear that we are not criticizing specialized language; word-concepts, individually and in discipline-specific collectives, are a necessary, associative structuring of knowledge. However, we *are* highlighting that the benefits of specialized language are matched by impediments, not the least of which is the inability for related bodies of knowledge represented by different sets of jargon to be easily revealed as related by simply keyword searches, e.g. strong/weak ties in sociology and interpersonal attraction in psychology. In addition, jargon changes within disciplines result in difficulties establishing ancestral-inheritance relationships in bodies of knowledge as they evolve over time, e.g. the Sapir-Whorf Hypothesis (Sapir, 1983) and linguistic determinism/linguistic relativism, or evolving definitions/conceptualizations of "culture" within anthropology. And also, the proveniences of seminal, inspirational thoughts to one discipline which originate from another discipline are difficult to reconstruct, e.g. Shannon's "A Mathematical Theory of Communication" (1948) to the theoretical underpinnings of the discipline of interpersonal communications. Such "cross-pollination" of disciplines may even be more difficult to achieve today given the scale of new and existing knowledge and knowledge sources, and the deepening of knowledge specialization. In addition, the presence of key words within a search result set of knowledge objects is not strongly predictive of the semantic context of the keywords thus guaranteeing a high noise-to-signal ratio in rough proportion to the scope and specialized/unspecialized quality of the knowledge bases being searched. "Noise", in this context, is a variety of conditions including result sets that are not relevant to intent and result sets which are impractically large (i.e. thousands or millions of results) regardless of relevance. These are just a few of the problems which simple keyword searches cannot resolve.

We assert that the simple keyword search is 1) kind of representation and structure that 2) increases complexity rather than reduces it by 3) being too "noisy", 4) dependent upon literal "tokens" of specific jargon, and 5) by flattening pre-existing structures into mono-dimensional set relationships without 6) sensitivity to meaning and context, thus 7) offloading the task to human beings of filtering and (re)relating knowledge unnecessarily, 8) increasing the difficulty of achieving "discovery", and 9) inhibiting the creation and publication of new, useful relationship between knowledge 10) in ways that can be recovered/discovered by others. By way of analogy, the simple keyword search creates a Flatland (Abbot, 1998) out of higher-dimensional knowledge-space.

We must also consider that keyword search result ranking in commercial search engines has a monetary value that forms the exchange-basis of the industry. That value is, in part, a result of the flaws of keyword searches

and predicated on the knowledge that there is a limit to how deep a given individual will search through the result set; that is, the manipulation of relevance has commercial value. Furthermore, this value-proposition extends to the commercial viability of specialized knowledge bases, such as databases of scholarly journal articles. Such knowledge bases limit the intrinsic flaws of keyword searches by offering searchable content that is presumed to be authoritative (because it is peer-reviewed) and discipline- or domain-specific. But even keyword searches within peer-reviewed databases may, and very often do, yield results of questionable relevance to the intent of the person searching. While the ability to access authoritative knowledge bases via the Internet represents a step forward, this advance is offset by 1) the need to either have preexisting knowledge of the existence of relevant knowledge bases or construct knowledge of the same, and 2) the resources to access these knowledge bases by either paying for access, or having an affiliation with an organization which pays for access. Keyword searches against the whole of the Internet treat all content as "equally" authoritative and relevant, after allowing for programmatic modifications to a result set for commercial gain.

The issues presented above are a significant challenge even for those with extensive research experience but, in the context of primary and secondary education, are potentially counter-productive by, for example, setting the conditions for the formation of negative attitudes about research, increasing the complexity of source evaluation and other forms of "critical thinking", forcing additional work in reconstructing "native" relationships among knowledge objects due to the decontextualization of keyword search result sets, and even inadvertent exposure to inappropriate content.

A search represents a learner's attempt to gain an entry point into a related web of knowledge. In both principle and practice, current keyword search technology creates unnecessary impediments and adds frustrations to this learning process. We believe that the problems intrinsic to keyword search technology may be mitigated by new search technologies such as correlation-based n-dimensional queries.

Signal: n-Dimensional Queries

We must move well beyond simple keyword searches as the predominant means of finding related knowledge objects. At stake is the premise that as our knowledge continues to grow, so too will the proportion of noise relative to signal we seek as the products of our intentionality. Searchable indices, even natural language full-text searches, are *efficient* when the individual conducting a search possesses a high degree of preexisting knowledge and a specificity of intention. But, as our knowledge continues to grow, we cannot assume that even experts will have the requisite degree of knowledge necessary for efficient keyword searches. Furthermore, keyword access only serves to reinforce lexical isolation between knowledge communities (and let us not forget that students also represent a kind of knowledge community); knowledge that is related conceptually but employing different discipline-specific words eludes the logic of keyword searches. But neither should we expect communities of knowledge specialists to abandon the invention of words or the reinvention of meaning to make our technologies work well; that would be placing the technological cart before the human horse.

Moreover, we must not only create new means to form knowledge objects into larger structures but also facilitate the sharing of these structures so that not just content but structure may be critiqued, tested, and refined. Often, the relationship between things is of equal or greater importance that the things themselves. Bush (1945) deftly stated this idea, "The inheritance from the master becomes, not only his additions to the world's record, but for his disciples the entire scaffolding by which they were erected." Embedded in this statement are, in fact, two scales and scopes of the problem: the problem the master encounters in finding, relating, creating, and assembling knowledge into the unique scaffolding that is the product of his or her expertise; and the means by which the master shares what he or she knows. For example, consider that a textbook represents not only a sampling of the canonical knowledge of a field but also a particular structural relationship between the objects of knowledge which make up that body. "History and systems" courses offer a more explicit view of socio-historical processes and the genealogies of ideas in the evolution of a discipline. Peer-reviewed journal articles impose internal structural relationships on content via the vector(s) of the narrative and create or afford external associative structures via citations, bibliographies, abstracts, and keyword lists. The Web affords structure via hyperlinks while journals create structure by containing knowledge related by the publisher's intent. Anthologies and conferences structure knowledge into conceptual themes. Computer applications such as Zotero and Endnote allow bibliographies to be shared. All of these structuring processes and tools – and the many others to which we have not referred – are necessary but not sufficient for the job at hand.

One of the most fundamental knowledge structures is the correlation. In this case, we are emphasizing correlation's meaning as a (statistically significant) association of one thing with another rather than the more

narrow meaning in the field of statistics of the degree with which two or more things are related or change together, although these two definitions are interdependent rather than exclusive. “Statistically significant” is parenthetically inserted because we are generally not interested in non-significant associations; these we include in the categories of “coincidence” or “noise” even in common experience. Correlation can be viewed as one of the most fundamental structures of learning, e.g. the probability-based associations formed between unconditioned stimulus and conditioned stimulus, and the importance of the strength of correlation for determining excitation verses inhibition in the Rescorla-Wagner Model (Rescorla & Wagner, 1972). Correlation may also been seen as the basis for much of scientific investigation; the pattern or anomaly in a pattern which inspires formal research. The importance here is that correlation appears to be a basic trait of human cognition and is therefore an excellent candidate for finding and structuring knowledge in “natural” ways. In addition, correlations are a beginning to structure, a "likely to be disconfirmed" (Kuhn, 1996) sort of structure and, thus, a natural laboratory for learners to engage higher-order critical thinking. Correlations are also ontological in nature; that is, a correlation pops into being as objects-phenomena relationship that must be at least temporarily reified to be investigated. The boogeyman of correlation is coincidence but, in ideal terms, even potential coincidence must be treated as if it is real for it to be falsified.

Searches based on correlations differ from keyword searches on a number of highly significant points. First is the requisite nature of the form of a correlation search: the need for two or more search terms rather than one. This is due to the nature of correlation itself – that is, the relationships between two things rather than set membership based on (at minimum) a single thing held in common. Correlation searches benefit from, but do not require, knowledge of specific jargon, a point we will elaborate upon more in a moment. Correlation search results are based upon analysis of “massive networks of associated ‘knowledge fragments’ to ‘connect the dots’ and establish relationships between terms, phrases, concepts, or topic” (Bobick & Cornwell, 2008) rather than index or index-like set association. Knowledge objects that make the most “valuable” connections to the network of correlations are considered to be the most relevant. Correlation searches are non-semantic – that is, correlation searches to not require the algorithmic determination of “meaning” in the search terms or the “corpus” (i.e. the knowledge based being searched) – but produce results which appear “intuitive” in quality. For example, a set of search terms such as “green tea”, “life quality”, “metastasize” can recover relevant documents related to cancer and polyphenols (a class of compounds that may help to combat cancer that are present in green tea). As this example, suggests, correlation searches can also reveal non-user-supplied (algorithmically determined) terms upon which the correlation results are based and thus become a resource to the searcher for specific jargon in a domain of knowledge. Correlation searches do <u>not</u> employ popularity or similarity metrics, binary logic, or link analysis. The import of this is that relationships revealed between knowledge objects are not necessarily predetermined by authors but, instead, are based on “native” or “organic” structure. While this feature of correlation searches risks decontextualizing knowledge objects in similar ways to keyword searches, the risk is at least partially offset by the ability to recover/discover previously unknown relationships between knowledge objects, and thus making correlation searches a potentially valuable tool for original research against the contents of the Internet or smaller, more focused knowledge bases. We refer to correlation-based searches as “n-dimensional” because each search term constitutes a dimension in an “answer space”. The precision of correlation searches increased substantially with the number of search terms, thus leading to smaller, more relevant – and more practicable – results sets.

Correlation-based searches will not replace keyword searches but do represent a significant and complementary advance over keyword searches. Nor do correlation-based searches represent the last word in search technology. What correlation-based searches do offer is the ability to at last employ a fundamental structure and method of knowledge against the burgeoning universe of human knowledge represented on the Internet. What else may be done with such a fundamental advance is anyone’s guess.

Conclusion: The Relevance to Education

The contents of the Internet represent the single largest collection of the human experience ever created and, therefore, in the *ideal* are an incalculably valuable resource for education. However, the Internet’s size and rate of growth alone presents daunting challenges to finding single items or small collections of relevant, useful materials. Without an intrinsic organizing principle beyond the technological layers of the network and the inherently limited functionality of keyword searches, the real-world value of the Internet to education is presently diminished. Research and development currently being conducted in other areas of the nexus between technology and education are important and valuable. However, we argue that for the Internet to achieve its potential in education there are fundamental issues with the structures of knowledge as represented on the Internet, and the means by which we find and utilize this knowledge, that must be addressed. In addition, we believe that such

fundamental advances will promote the further development of higher-order projects like those to be presented at this conference.

We submit that correlation-based search technology represents this kind of fundamental advance because it addresses many of the confounding limitations of keyword searches as the dominant paradigm for "selection". First and foremost is helping learners to more efficiently discover entry points into relevant webs of knowledge, without the degree of preexisting knowledge required by even the most advanced keyword searches. Without the additional burdens presently imposed on learners because of the limited affordances of available search technologies, learners still face great challenges in assimilating, evaluating, and communicating new knowledge which they may find, skills which we might all agree are of greater importance than the act of searching.

References

Abbott, E. A. (1998). Flatland : a romance of many dimensions. New York: Penguin Books.

Barnes, K., Marateo, R. C., & Ferris, S. P. (2007). Teaching and learning with the Net Generation [Online]. *Innovate Journal of Online Education, 3*(4).

Bobick, M. & Cornwell, W. R. (2008). A Knowledge Correlation Search Engine: Technical White Paper. The Center for Internet Research. Retrieved December 9, 2008, from http://tcfir.org/whitepapers/A%20Knowledge%20Correlation%20Search%20Engine.pdf.

Bush, V. (1945). As we may think. *Atlantic Monthly*. Retrieved August 15, 2008, from http://www.theatlantic.com/doc/194507/bush.

Kuhn, T. S. (1996). The structure of scientific revolutions (3rd ed.). Chicago, IL: University of Chicago Press.

Lyman, P., & Varian, H. R. (2000). How much information 2003? *School of Information Management and Systems, University of California at Berkeley*. Retrieved August 12, 2008, from http://www2.sims.berkeley.edu/research/projects/how-much-info/

Lyman, P., & Varian, H. R. (2003). How much information 2003? *School of Information Management and Systems, University of California at Berkeley*. Retrieved August 12, 2008, from http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/

Rescorla, R. A., and Wagner, A. R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement, *Classical Conditioning II*, A. H. Black and W. F. Prokasy, Eds., pp. 64-99. Appleton-Century-Crofts.

Sapir, E. (1983). Selected writing of Edward Sapir in language, culture, and personality: University of California Press.

Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27, 379-423.

Tapscott, D. (1998). *Growing Up Digital: The Rise of the Net Generation*. New York: McGraw-Hill.