

Textual Behavior in the Human Male

Stephen Ramsay

March 30, 2012

My title, of course, refers to the infamous Kinsey Report: *Sexual Behavior in the Human Male*, first published in 1948.

[SLIDE]

The Kinsey report is, like *Harry Potter*, *The Omnivore's Dilemma*, and *The Book of Job*, one of those books you feel you've read even if you haven't actually read it. Its general outlines and conclusions, and the national sense of scandal that ensued, are well known even after sixty years. When you do sit down to read it, though, the most shocking thing isn't its prurience or its candor, but its easy, approachable tone. It sounds like this:

The English people are more or less justly reputed to be the most completely clothed people in the world, and Americans have been slow in breaking with English tradition. The American visitor to foreign lands is often amazed at the exposure which is allowed in some other cultures, and he criticizes it on moral grounds. The nudity of the French burlesque is ascribed to the "low morality" of the Frenchmen as a group; and although an approach is made to the same sort of display in American burlesque, the institution here does not achieve the same free acceptance of complete nudity which the original French has. (365)

On kissing:

...The lower level male [Kinsey is speaking here of lower socio-economic class] considers [deep kissing] to be dirty, filthy, and a

source of disease, although he may drink from a common cup which hangs on the water pail, and he may utilize common utensils in eating and drinking. Obviously, the arguments, at both levels, have nothing to do with the real issues. They are rationalizations of mores which place taboos upon mouth contacts for reasons which only the student of custom can explain. (369)

This tone—learned, somewhat detached, but entirely approachable and plain-spoken—is the same tone assumed by Max Weber, Vannevar Bush, Claude Lévi-Strauss, Émile Durkheim, Walter Ong, Clifford Geertz, and even, in his less strident moments, Marshall McLuhan. No self-respecting sociologist would write like this today.

[SLIDE]

Yet all of this almost-belletristic prose is accompanied by hundreds of earnest-looking graphs and tables that mark it, unmistakably, as of a piece with modern science. You have to wade through a full hundred-and-fifty pages of detailed methodological discussion to get to the good parts, and even when you get there, the text wavers back and forth between cold data and hot description. One minute, he sounds like Lewis Mumford; the next minute, he sounds like (or rather, looks like) a man delivering a report on the hydrology of the midwest. But make no mistake, Kinsey seems to say: all of this is science. And that means something very particular. Here's Kinsey again, in the introduction to the work:

The present study...represents an attempt to accumulate an objectively determined body of fact about sex which strictly avoids social or moral interpretations of the fact. Each person who reads this report will want to make interpretations in accordance with his understanding of moral values and social significances; but that is not part of the scientific method and, indeed, scientists have no special capabilities for making such evaluations [...] This is first of all a report on what people do, which raises no question of what they should do, or what kinds of people do it. It is the story of the sexual behavior of the human male, as we find him. (5;7)

The extraordinary naiveté of this paragraph is entirely patent to any humanist. The human male “as we find him?” Each person “will want to make interpretations?” It's scarcely necessary to rehearse the problems with such ideas, and for

the record, absolutely no one bought it. In fact, the Kinsey report bears the Library of Congress subject heading “United States—Moral Conditions.”

But Kinsey’s insistence that this is “just the facts” is not simply a moment of wishful thinking—a dodge intended to tamp down the raging criticisms that he knew would come, and which did come from all quarters. It belongs, rather, to a moment in history that starts around the turn of the century, continues until just after the Second World War, and which arguably occurs in more sublimated forms up until the present day. What Kinsey is struggling with, is the scientization of the humanities.

[SLIDE]

Sex and the erotic was, after all, for most of the history of the west, *our* department—which is to say, the province art and of humanistic inquiry. Freud barely even attempts to scientize his subject matter. A great number of Freud’s basic ideas form the ground truths of modern psychology, but his hermeneutical framework lives on mainly in English departments. Vladimir Nabokov’s legendary quip—“Let the credulous and the vulgar continue to believe that all mental woes can be cured by a daily application of old Greek myths to their private parts” (130)—is not far off the mark, inasmuch as Freud’s basic method is fully embedded in the traditions of literary criticism. The same might be said of Marx, whose most scientific source text is probably the *Nichomachean Ethics*.

But Kinsey appears at a time when nearly all the subjects of humanistic concern are being reexamined as questions susceptible to scientific methodology. What is the meaning of work? What is the imagination? How do I live a good and fruitful life? What is love? Ian F. McNeely and Lisa Wolverson, in their book, *Reinventing Knowledge*, locate the rise of the social sciences in the twentieth century precisely in the attempt to explore such questions using scientific methods:

The ability of modern science to manipulate and master nature, to command public esteem, and to change public behavior arose from the synergy of a craft workshop fused with a disciplinary seminar. [...] Social scientists rose to influence by applying laboratory techniques to places where people now learned, worked, and lived. Disciplines such as economics, sociology, and anthropology were already developing around traditional philological methods. Writers like Adam Smith, Max Weber, and Émile Durkheim provided a body of canonical texts for endless analysis, and written sources like gov-

ernment statistics and travel reports gave their acolytes the means to produce new scholarship. [...] Social scientists became intelligence testers, efficiency experts, scientific philanthropists, and much else besides. Humanists transformed into scientists, they took their white-coated counterparts as a model. They measured and quantified; they indoctrinated cadres of experts in experimental methods; and they crossed into the public domain bearing “objective,” impartial findings meant to effect widespread social change. (229-230)

The catalog of results, even in the less obviously appalling instances—Adolf Hitler was exceedingly fond of such methods—is the sort of thing that makes humanists cringe.

[SLIDE]

Standardized testing began with these early forays into scientized humanism, with the result that “Every human being, whatever his or her particular constellation of intellectual faults and fortes, carries a two- or three-digit ‘intelligence quotient’ (IQ), largely invariant over a lifetime, that denotes his or her inherent mental capacity” (230).

[SLIDE]

Students of Taylorism, armed with stopwatch, clipboard, and camera, attempted to break every factory worker’s movements into their constituent parts in search of efficiency. Statistics became the principal tool wielded by large philanthropic institutions, so that attempts to cool race relations and eradicate poverty could be based not on traditional dialectical methods and religious philosophies, but on cold hard facts. The Kinsey report itself, was funded mainly with Rockefeller money.

[SLIDE]

We cringe, not because we dispute the facts that more modern instantiations of these methods provide, but because we are suspicious of the ability of facts to translate into answers to questions that have persisted—literally, for centuries—within our disciplines. The integrity of Kinsey’s work as science was sharply criticized almost immediately. Yet in the end, Kinsey’s conclusions—much like

Freud's—still stand. Many more people masturbate than say they do; the range of sexual experience differs according to both race and class; homosexual and heterosexual (terms which Kinsey studiously avoided as ontic descriptors) describe a broad spectrum of sexual behaviors and feelings experienced by most humans over the course of their lives. These were, in Kinsey's estimation, the facts of the case. If he cannot hold himself to dispassionate dilation of those facts, it is not because such a project is impossible, but rather because there was no available discourse for relating scientific facts to humanistic conclusions. The ancient pornographer knew precisely "What he actually does." But what does it mean?

There are several varieties of digital humanities that offend no one. The student of human culture can hardly object to the work of those who have labored tirelessly, these last few decades, to bring the written artifacts of human experience online. The study of digital culture itself—media studies and its ramified offshoots—is likewise so consistent with the ordinary hermeneutical methods of the humanities as to be barely noticeable as a "new thing." Even those suspicious of work in educational technology will concede the general utility of such research; the fact that Blackboard sucks is not usually meant to condemn the entire project of bringing technology to bear on teaching.

[SLIDES]

What strikes fear in our hearts, is the work of those use computers in order to engage in the quantitative analysis of cultural artifacts. People, to put plainly, like me.

My own career development as a digital humanist is very much like what McNeely and Wolverson describe. Like the early social scientists, I was trained entirely in the art of the seminar and in the deployment of theoretical frameworks, which, though seeming very new at the time, were methodologically descended from the dialectical and philological traditions that began during the rise of the modern university in nineteenth-century France and Germany. Becoming a digital humanist meant—and still means—learning to operate within a laboratory environment. A digital center today is, in fact, very much like a chemistry or biology lab. Knowledge is conveyed not primarily through free-flowing, seminar-style discussion, but through collective and collaborative mastery of complicated equipment and difficult methodologies. "Thing knowledge" is essential.

We encounter, too, some of the objections that were leveled at nineteenth-century science, which (we should recall) had great difficulty inserting itself into the university. In telling the story of modern science, we are apt to focus on

resistance to the work of scientists like Darwin, whose conclusions (much like Kinsey's) challenged numerous assumptions that had been in place for centuries. But the more immediate problem for the likes of Pasteur or Lavoisier is that their work looked for all the world like plumbing. Laboratories, for most of the nineteenth century, were housed in peoples' homes, and relied heavily on developing the skills necessary for fabricating and manipulating equipment for the purpose of making accurate measurements—something that had always been a non-academic craft skill. From the standpoint of the Classics faculty at Paris, nothing could be more base. Even the words Lavoisier used to describe his work—"carbonate," "nitrate," "sulfate"—seemed barbaric to the French intelligentsia (McNeely 217), and presumably rolled off their tongues much as the jargon of computation does today for many members of the humanities faculty.

And what is it, precisely, that we digital humanists are fabricating? Susan Hockey, one of the pioneers of digital humanities in its modern form, offered a definition of our principal toolset that is as uncontroversial as it is dispiriting:

Computers can assist with the study of literature in a variety of ways, some more successful than others. [...] Computer-based tools are especially good for comparative work, and here some simple statistical tools can help to reinforce the interpretation of the material. These studies are particularly suitable for testing hypotheses or for verifying intuition. They can provide concrete evidence to support or refute hypotheses or interpretations which have in the past been based on human reading and the somewhat serendipitous noting of interesting features. (66)

The text as we find it—an objectively determined body of fact about text which strictly avoids social or moral interpretations of the text. Each person who reads our work will want to make interpretations in accordance with his or her understanding of moral values and social significances; but that is not part of the digital humanities and, indeed, digital humanists have no special capabilities for making such evaluations.

The genealogy I have been tracing suggests a clear teleology for digital humanities. Digital humanities, which, in its data-centric forms, represents a set of practices that centers on tool building and instrumental methods of analysis, represents yet another attempt to scientize the humanities. It will struggle mightily to insert itself fully into the academy, but once it does, it will make a clean break with its more purely contemplative past. The social sciences are not "the

sciences;” *digital* humanities will not be the humanities. A few months ago, I was at a conference on data mining in the humanities. During one of the breaks, a quite prominent historian—well known for his use of quantitative methods—leaned over to me and averred, “I used to be an historian. Now I’m a social scientist.”

The question I’d like to propose today is simply this: How do we prevent this from happening?

Some, it should be said, would very much like this to happen, and see digital humanities as the way forward. Jonathan Gottschall, in a 2008 editorial in the *Boston Globe*, described the field of literary studies as “moribund, aimless, and increasingly irrelevant to the concerns not only of the ‘outside world,’ but also to the world inside the ivory tower.” The solution is one that even C. P. Snow would have found provocative:

I think there is a clear solution to this problem. Literary studies should become more like the sciences. Literature professors should apply science’s research methods, its theories, its statistical tools, and its insistence on hypothesis and proof. Instead of philosophical despair about the possibility of knowledge, they should embrace science’s spirit of intellectual optimism. If they do, literary studies can be transformed into a discipline in which real understanding of literature and the human experience builds up along with all of the words.

This proposal may distress many of my colleagues, who may worry that adopting scientific methods would reduce literary study to a branch of the sciences. But if we are wise, we can admit that the sciences are doing many things better than we are, and gain from studying their successes, without abandoning the things that make literature special.

I, for one, react to this idea with horror. I’ll take philosophical despair any day of the week.

Let me say that I am precisely the sort of the person who should be clamoring for this mystical singularity in which the humanities sloughs off its ancient proclivity for free-flowing discussion of questions without answers and embraces the optimistic joy of scientific positivism. I have spent my whole career as an English professor studying computers. I am proficient in at least a dozen programming languages; I have learned the arcana of statistical mathematics; I can hold my own in a conversation with any computer scientist; and I spend most of my days either building tools or teaching unsuspecting English and History majors how to

do the same (sometimes in the Department of Computer Science). What's more, my methods are more-or-less exactly like those of Gottschall, whose work is fascinating and focuses mainly on computational analysis of text. I am an enthusiastic and active member of the revolution.

But my own sense of what makes the humanities "special" is its alternative ways of knowing and talking. I have nothing but the highest respect for scientific ways of knowing and talking, social or otherwise. But a world in which the mysteries of the human condition are unfolded without the methodologies of the humanities seems to me a poorer world. So when I say, "How do we prevent this from happening," I really mean: How do we move forward with the wondrous tools and methods of digital technology while still remaining faithful to ourselves?

Because the tools are, indeed, wondrous.

[SLIDE]

Here's a list of English novels created by David Hoover, a professor of English at NYU, a long-time member of the digital humanities community, a brilliant text analysis scholar, and someone with whom I disagree on a quite regular basis. The list was produced by writing some code that orders texts in terms of "vocabulary richness," which Hoover defined, for the purposes of this experiment, as the number of different words per 50,000-word block. That is to say, it organizes novels in terms of how wide the vocabulary is in a particular author.

I've shown this list to lots of people over the last few years, including quite a few professors who are familiar with every novel on the list. Most have an easy time placing either Henry James or William Faulkner at the top of this list, perhaps alongside either Mark Twain or Oscar Wilde. *Kim* and *The Sea Wolf*, both books that are frequently offered to young adult readers, presumably belong somewhere near the bottom. Perhaps you are thinking of this slightly differently, with well-conceived—if only briefly considered—arguments for where the books belong on the spectrum of vocabulary richness.

But the truth is that no one I've ever showed it to—including the English professors—have gotten it remotely correct. This list, you see, is in *ascending* order of vocabulary richness. Faulkner and James at the bottom in terms of rich vocabulary, Kipling and Lewis at the top.

And then it starts. The first objection is to the notion of "richness," which, you'll recall, is clearly stipulated when the problem is presented. Most will complain that "richness" has nothing at all to do with the number of different words, though when I ask what "richness" is, it's as if I asked them a question about the

nature of the Trinity. Others react by quickly explaining to me why these results are clearly correct, despite having just given me the wrong answer. Such explanations can be quite elaborate, very well thought out, and entirely enlightening, but they're delivered in a tone that says, "Yes, yes, of course. I knew that." Other text analysis practitioners will immediately begin talking about sample size, tokenization, and other technical aspects of the methodology (as if to say, "I would have known that, if you had conducted the experiment properly.")

The disposition of the researcher is instructive as well. Hoover, who created the list, is solidly of the positivist party. In an article entitled, "The End of the Irrelevant Text: Electronic Texts, Linguistics, and Literary Theory," he writes:

Much high theory is deeply influenced by ideas about the instability of the sign and the tendency of texts to disintegrate under critical pressure, ideas most closely associated with the late Jacques Derrida. [...] Jerome McGann, for example, champions the game-like, "fundamentally subjective character of . . . criticism" [McGann 2004, 50–1] and asks "What if the question isn't 'how could he [the critic] take himself or his ideas seriously' but 'why should he take himself or his ideas seriously'?" [McGann 2004, 50]. Stanley Fish emphasizes the reader's role, attacking the idea that texts have meaning at all—arguing that the only links that exist between the text and its interpretation are those that are "fashioned in response to the demands of the reading experience" [Fish 1980, 64]. Critical approaches like these and a more general distrust of the sign/signified link within literary theory helped to produce a climate inhospitable to text-analysis and stylistics.

Hoover's project, briefly, is to destroy this arrant nonsense with the power of technology. You thought you knew what was going on? Well, you don't. I will admit to a certain guilty pleasure in showing people this list; it is a bit of a gotcha after all. I usually don't mention my own reaction when the experiment was conducted on me: astonishment, followed by the irresistible urge to go try it on someone else.

But it's the contours of the resulting conversation, and our sense of what it is we're doing with this kind of work, that fascinates me—in particular, the ways in which reactions to Hoover's list closely resemble reaction to the Kinsey report. We want to talk about methodology, or we want to talk about how we knew it all along. We want to stand back and present results as "the facts," or we want

to critique the notion of fact itself. We become, in other words, something like humanists transformed into scientists, taking our white-coated counterparts as a model. What is missing from such conversations, it seems to me, is actual humanistic discussion.

Let's begin by admitting that Hoover's list is probably correct as an experimental observation. It might be flawed in dozens of ways, but I suspect it isn't, and no matter how many experiments we conduct, we're eventually going to have to accept or reject conclusions that are, like both scientific and humanistic conclusions, largely the result of inductive reasoning. Let us also admit that these results surprise us; it's just not the way we think of these books. Then, instead of trying to explain it all away as knowledge that is somehow already fully tractable within our current paradigms, let us ask, "What on earth does this mean?"

It might suggest peculiar mappings between the way books are understood in general as cultural objects and the way that comes to influence the way they are experienced by individual readers. It might show us the way to a deeper understanding of reading itself and of complexity as something that is experienced conceptually, but then thought to inhere in textuality in some more literal manner. It raises any number of questions about the evolution of English style, which is only dimly understood even by experienced linguists. Most importantly, it suggests study of our own discursive practices—not only interrogation of our critical paradigms, but further awareness of the way we conduct discussion within those paradigms.

Over the years, my work has become a quest for meaningful astonishment, and I had a particularly stirring moment a few years ago while working on Virginia Woolf's 1931 novel *The Waves*.

[SLIDE]

[The Waves] consists of a series of monologues that trace the lives of six friends from early childhood to old age, each monologue (beginning always with "Susan said" or "Bernard said") telling the characters' stories at seven distinct stages of their lives. Yet "story" is far too strong a word for their ruminations. The characters recount only a few of the sorts of events one would expect to see forming the basis of plot in a conventional narrative. They speak about different things and have different perspectives on the world, but they all speak in roughly the same manner, and do so from childhood

to adulthood—employing, as one critic puts it, “the same kind of sentence rhythms and similar kinds of image patterns” throughout (Rosenthal 144). Some critics have suggested that there are differences that lie along the axis of gender or along a rift separating the more social characters from the more solitary ones, but in the end, one has the sense of an overall unity running against the perspectival conceit that frames the narrative. (Ramsay 9-10)

This is precisely the sort of problem that attracts text analysis practitioners. So a graduate student of mine (Sara Steger, now at the University of Georgia) and I began running classification and clustering algorithms on Woolf’s text, precisely to see if there were similitudes that are hard to arrive at using, to borrow Hockey’s phrase, “ordinary reading.” The work was going well, mostly because we spent most of our time trying to understand what it was we were looking at. On a lark, we decided to pose a simple question: What are the words that the women in the novel use, but which none of the men use? Here’s what we found:

[SLIDE]

We then did the same for the men:

[SLIDE]

Did you just gasp? So did we. Some critics find this list terribly upsetting—convinced that I’ve done something wrong. Others feel absolutely confirmed by it (the sexism of the Western canon, pace Gottschall, writ large). But if you’re anything like the many scholars I’ve shown this to over the years, your immediate reaction is to resolve either disposition with further experimentation. Perhaps the women are more isolated? Or their vocabulary is more rich (how do we do that richness thing again)? Were I to hold the Q&A right this minute, I suspect I’d have a half-a-dozen hands in the air with a question that begins, “Have you tried...?” Have you tried it with other novels, other time periods, other character breakdowns, by author gender, by genre, with Jane Austen? If you’re thinking this way, perhaps you’ve caught the DH bug. It is my pleasure to welcome you.

But I must ask: What is it that we are trying to know? Are we trying to find an *answer*? And if so, what is the question? Do we imagine that such further experiments would resolve long-standing questions about gender and language? Do we really want those questions resolved?

That last question may seem slightly perverse, but I believe that, in the end, what is most distinct about humanistic inquiry is its resistance toward final answers. It is the goal of the seminar to answer questions, but mostly by proposing them more fruitfully. The humanities wants for itself a world that is more complex than we thought—a sense of the human experience as deeper and more surprising. We are in search of a conversation, really, and the thing that has always sustained that conversation is the artifacts of the human experience (especially, written artifacts).

In 2015, the Text Creation Partnership will release into the public domain nearly every title from the first two centuries of English print culture. *The Perseus Project* right now contains nearly the entire extant corpus of ancient Greek. The digital Library of Latin Texts contains nearly every important work written in Latin between 240 BCE and the Second Vatican Council. Google Books is now rapidly approaching seventeen million volumes. I could go on. These resources are so new, that nearly every question we ask of them is one that has never been asked before—largely because without digital tractability, our only real option was to read them all (an impossible task, by any measure).

It is customary to say that such wonders will transform the humanities, but that is not at all clear to me. The technical requirements for the study of large-scale corpora are often quite significant (though I should mention that my students are writing programs like Hoover's and mine with a few weeks), but that is not even the limiting factor. The limiting factor, it seems to me, is how (and, perhaps, whether) we decide to remain humanists in the face of such abundance. We can do what we've always done; we're interested in talking about books, and now we will read them online. The other option, though, seems to me the more exciting one: learning to read anew using the tools of the new technology. At least some of us need to become programmers, tool builders, text engineers. But many more of us need to decide to invite the new texts that result from that effort—frequency lists, visualizations, n-grams, maps, data mining results, and much more—into our ongoing conversation.

That might mean, ironically, learning to write like Kinsey—oscillating back-and-forth between essay and data, *belle lettres* and *numerique*—without the pretense of scientific positivism. Perhaps then we can create works that attempt to accumulate a subjective body of rumination about text which strictly embraces social or moral interpretations of the text. Each person who reads such a report will want to make interpretations in accordance with his or her understanding. That is part of the humanistic method, and indeed, humanists have unique capabilities in making such evaluations.

[SLIDE]

Works Cited

- Appel, Alfred, Jr. "An Interview with Vladimir Nabokov." *Wisconsin Studies in Contemporary Literature*. 8.2 (1967): 127–152.
- Gottschall, Jonathan. "Measure for Measure." *The Boston Globe* 11 May 2008, third ed.: D1.
- Kinsey, Alfred C., Wardell B. Pomeroy and Clyde E. Martin. *Sexual Behavior in the Human Male*. Philadelphia: Saunders, 1948.
- McNeely, Ian F. and Lisa Wolverson. *Reinventing Knowledge: From Alexandria to the Internet*. New York: Norton, 2008.
- Ramsay, Stephen. *Reading Machines: Toward an Algorithmic Criticism*. Urbana: U of Illinois P, 2011.