

Contexts

<http://ctx.sagepub.com/>

An Old Tool with New Promise

Andrew M. Lindner

Contexts 2012 11: 70

DOI: 10.1177/1536504212436505

The online version of this article can be found at:

<http://ctx.sagepub.com/content/11/1/70>

Published by:



<http://www.sagepublications.com>

On behalf of:



Additional services and information for *Contexts* can be found at:

Email Alerts: <http://ctx.sagepub.com/cgi/alerts>

Subscriptions: <http://ctx.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Feb 14, 2012

[What is This?](#)

an old tool with new promise

by andrew m. lindner

Even with two ongoing wars, unrest in the Middle East, and historically high rates of unemployment, in June 2011, news coverage of Congressman Anthony Weiner's sexually explicit messages and photos accounted for a full 17 percent of the total national print space and radio and TV airtime available for news. Since late 2006, Pew Research Center's Project for Excellence in Journalism (PEJ) has conducted weekly content analyses of national news programming. This ongoing assessment of news programming provides us with powerful insights into the national media and the kind of information it circulates. However, until fairly recently, this type of content analysis would have seemed wildly

short-lived decline (see top right). Since the early 2000s, however, content analysis has become an increasingly popular research tool. Beyond sociology, the number of social science, arts, and humanities articles mentioning "content analysis" and catalogued by Google Scholar grew dramatically in the 2000s, both in terms of raw numbers and as a percentage of all academic work (see bottom right).

The instant accessibility of content of all types has dramatically reduced the labor entailed in collecting data. In 1900, Delos F. Wilcox, a political scientist and lifelong fruit farmer, published one of the earliest modern quantitative content analyses, counting the number of columns dedicated to sixteen different

shows almost instantaneously.

In 2004, a team of researchers led by sociologists Doug McAdam, John McCarthy, Susan Olzak, and Sarah Soule coded *New York Times* coverage of protest events since 1965 to examine a range of questions about social movements. This mammoth data collection project became possible when the *Times* made all of its articles since 1851 available electronically online. Using complex search terms, the research group was able to identify and download all articles on protest within a specific period of time—a project that might have taken months or years to complete in a physical archive.

Today, researchers can use a variety of text analysis software packages to expedite and automate the process of coding, allowing them to search for the frequency of a specific word (say, "nurture") or a group of words (nurture or cuddle or snuggle) or to test whether a group of words is more common in one group of texts than another ("Do parenting books aimed at women use significantly more intimate words like nurture than those aimed at men?"). Word clouds, which use a very simple form of this technology to visually represent the frequency of words in a text, are particularly popular online. For example, *Washington Post* blogger Ezra Klein's "word cloud" of President Obama's 2011 State of the Union address revealed that the words "new" and "jobs" loomed large.

More advanced quantitative programs have their own dictionaries, which can tell researchers, for example, how

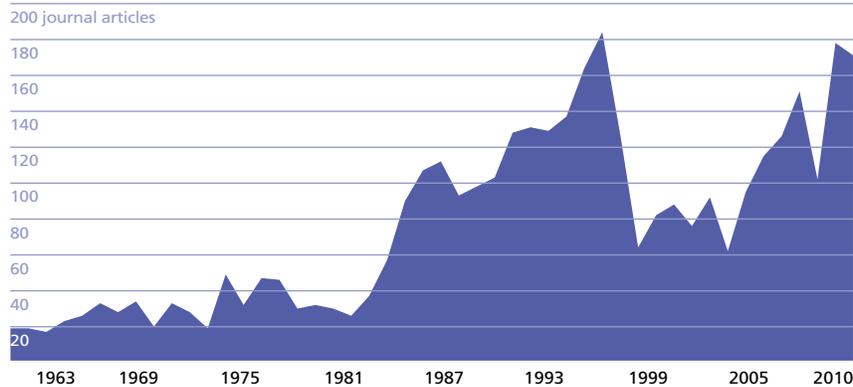
Content analyses offer partial insights into human behavior, but they can't tell us why people protest, why law firms hire certain lawyers, or why football coaches do what they do.

ambitious—maybe even impossible—to produce.

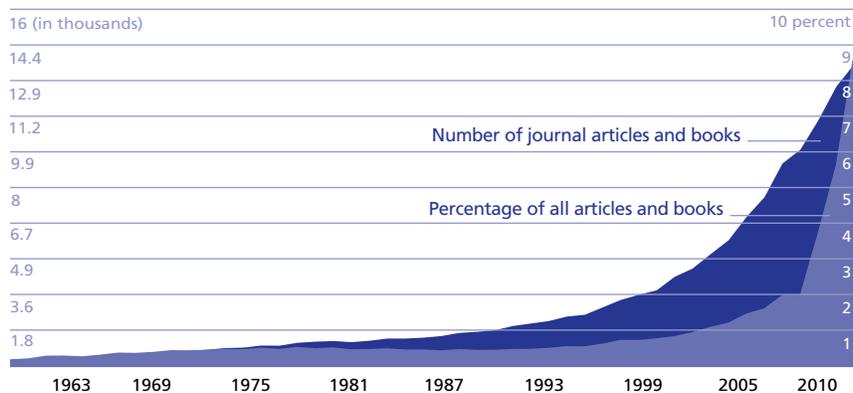
Content analysis is not new; sociology journals of the 1950s were filled with "propaganda analyses." But over the past 20 years, the instant availability of hundreds of television shows, decades of news reports, and millions of pages of official documents in electronic form have enabled faster and more cost effective content analyses. The explosive growth of content analysis in sociology in 1990s was followed by a precipitous, if

types of content (including war, political crime, and other issues) in 147 American newspapers. More than a hundred years later, the mission of Pew's weekly content analysis of national news programming is similar. What is profoundly different, though, is the manner in which this information is collected. Poor Wilcox had to collect physical copies of far-flung newspapers and add up the results by hand. Today, PEJ is able to access electronic copies of newspapers, television news programs, and radio

Timeline of *Sociological Abstracts* content analysis articles



Timeline of content analysis articles* on Google Scholar



*Social Sciences, Arts, and Humanities database only

often the content indicates traits such as activity, certainty, or optimism. Of course, some researchers choose to construct their own dictionaries and grammars to better suit the needs of their project. Many programs can also create visualizations, like cluster analyses, which show which words are most likely to appear together in content and be used in similar ways. For a sociologist interested in media depictions of Muslims, for example, this tool might be particularly helpful in figuring out not only whether the words “Muslim” and “terrorist” both appear frequently in transcripts from cable news shows, but also whether the two tend to appear together.

While increasing the efficiency of content analysis, these innovations raise new concerns about the meaning of findings in an age of limitless information and automated coding. Computers—though deploying sophisticated

software—lack the intellectual ability to intuit and interpret. Software can identify instances of “Muslim” and “terrorist” appearing together, but can’t tell us about the essential meaning of the sentence or the intent of the author. Although generating intriguing findings is now easier, the challenge is figuring

Content analysis generates intriguing findings—but the challenge of finding out what they really mean remains.

out what they really mean.

Creative new uses of the method are also permitting social science researchers to study “real world” phenomena and glimpse spaces off-limits to researchers—such as the hiring room at a large law firm. Sociologist Elizabeth Gorman studied gender bias in hiring

practices at such firms by linking content analyses of job advertisements with data about who was ultimately hired for the positions. She coded whether the job ads used stereotypically masculine (ambitious, assertive, quantitatively-oriented) or feminine (cooperative, friendly, verbally-oriented) criteria. She also recorded the hiring decision-maker’s gender, the proportion of women among the firm’s associates, and the firm’s areas of practice. She then used the law directory from the next year to code the gender of the hired person based on her or his name.

As she suspected, Gorman found that firms using more stereotypically masculine and fewer feminine terms in their job advertisements were less likely to hire female lawyers. A firm’s gendered understanding of who would make a good legal associate, in other words, affects their hiring practices. (Of course, these findings tell us nothing about how the lawyers making hiring decisions actually *think*.)

In a very different social context, political economist David Romer uses the indirect lens of content analysis to

examine decision-making within the NFL. He examined coaches’ decisions during fourth downs in the first quarter to kick a field goal or to try to advance the ball ten yards for a first down. Using play-by-play transcripts from over 700 games, he coded information on the yard-line, the decision, and the play’s

trends

outcome, establishing whether teams were, on average, better off kicking or going for a first down in a variety of situations. He then compared the ideal decision in a given situation to the coach's actual decision.

Despite the popular view that football coaches are chieftains of rationality, it turns out their judgment wasn't very good. Deeply conservative, the coaches generally choose the safe decision to kick rather than trying for a first down that might lead to a touchdown. In situations in which kicking would be the optimal choice, coaches usually made the right decision, but, as Romer notes, "Of the 1,068 fourth downs for which the analysis implies that teams are on average better off going for it, they kicked 959 times."

In another example, looking at the

motivations of suicide bombers during the Second Intifada using reports from newspapers and government and policy center web sites, sociologists Robert J. Brym and Bader Araj analyzed the individual motives, organizational rationales and events that preceded 138 suicide bombings by Palestinians. Contrary to past research, which had emphasized the strategic nature of such decision-making, Brym and Araj found that 82 percent of attacks were actually reactive, motivated by a desire for revenge against Israeli actions.

That they could come to such conclusions about the motives of suicide bombers using information only available online speaks to both the strengths and risks of content analysis. While Brym and Araj's analysis allowed them to systematically document reasons for suicide

bombings without putting themselves in danger, it fails to capture the complex motivations and personal details of the suicide bombers' lives—something that ethnography, for example, might reveal.

In other words, content analyses offer partial insights into human behavior. They can't tell us why people protest, why law firms hire certain lawyers, or why football coaches do what they do; the method doesn't give research subjects a voice. Still, as masses of new information come online, content analysis will continue to offer researchers ways of developing powerful insights into the social world.

Andrew M. Lindner is in the sociology department at Concordia College in Moorhead, Minn. He studies politics, stratification, and culture.

america's lost generation

by sangyoub park

Those who came of age during Japan's economic tumult of the 1990s became a generation of discouraged and reclusive people. Much of what Japan endured, including volatile real estate markets and high unemployment, is occurring in the U.S. today. If the recession continues to linger, America could be creating its own "lost generation," characterized by dissatisfaction and insecurity.

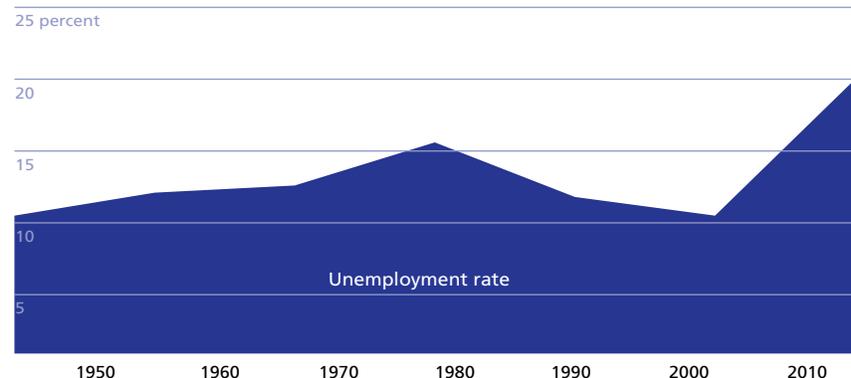
"Generation," one of the most widely used concepts in the social sciences, refers to an aggregate of individuals born around the same time who experienced some socially significant events in common. These shared, formative experiences frequently contribute to a unique worldview and remain powerful throughout individuals' lives. Since each generation has its own zeitgeist, the concept of generation is a useful tool for exploring social differences.

The Great Recession, though officially over, is nonetheless affecting every

aspect of Americans' lives—especially those of young people. Rates of youth unemployment drastically increased from 9.6 percent in 2000 to 19.1 in 2010, while national unemployment rates hovered between nine and ten percent during the past two years. (Figure below illustrates changes in the youth unemployment rate between 1950 and

2010.) Specifically, unemployment rates for young men (20.5 percent in 2010) are higher than those for young women (17.5 percent in 2010). Unemployment rates for young black Americans soared to 33.4 percent in 2010 (see top right). With unprecedented unemployment rates, there is little evidence that America's economy is improving.

Unemployment Rate, 16-to-24-year-olds, 1950 to 2010



Source: Bureau of Labor Statistics