Colleague Crowdsourcing: A Method for Fostering National Student Engagement and Large-N Data Collection

Amber E. Boydstun, *University of California, Davis*
Jessica T. Feezell, *University of New Mexico*
Rebecca A. Glazier, *University of Arkansas, Little Rock*
Timothy P. Jurka, *University of California, Davis*
Matthew T. Pietryka, *Florida State University*

ABSTRACT Scholars often rely on student samples from their own campuses to study political behavior, but some studies require larger and more diverse samples than any single campus can provide. In our case, we wanted to study the real-time effects of presidential debates on individual-level attitudes, and we sought a large sample with diversity across covariates such as ideology and race. To address this challenge, we recruited college students across the country through a process we call "colleague crowdsourcing." As an incentive for colleagues to encourage their students to participate, we offered teaching resources and next-day data summaries. Crowdsourcing provided data from a larger and more diverse sample than would be possible using a standard, single-campus subject pool. Furthermore, this approach provided classroom resources for faculty and opportunities for active learning. We present colleague crowdsourcing as a possible model for future research and offer suggestions for application in varying contexts.

Much of our discipline's understanding of political attitudes and behavior has been developed through studying two common groups: nationally representative samples and college students. Nationally representative samples are expensive and often lack internal validity; however, by design, they have high external validity. Student samples, although less representative, are often less expensive and can better facilitate experimental designs, providing strong internal validity. In this article, we present colleague crowdsourcing as a complementary research design that leverages strengths of each approach, and we illustrate its worth in a study of presidential-debate effects. We find that crowdsourcing not only facilitated our data collection but also engaged many students in active learning about the debates in ways that they otherwise might not have experienced. Thus, colleague crowdsourcing has benefits for both research and teaching.

Amber E. Boydstun *is an assistant professor in the political science department at the University of California, Davis. She can be reached at aboydstun@ucdavis.edu.*
Jessica T. Feezell *is an assistant professor in the political science department at the University of New Mexico. She can be reached at jfeezell@unm.edu.*
Rebecca A. Glazier *is an assistant professor in the political science department at the University of Arkansas, Little Rock. She can be reached at raglazier@ualr.edu.*
Timothy P. Jurka *is a graduate student in the political science department at the University of California, Davis. He can be reached at tpjurka@ucdavis.edu.*
Matthew T. Pietryka *is an assistant professor in the political science department at Florida State University. He can be reached at mpietryka@fsu.edu.*

COLLECTING DIVERSE LARGE-N DATA IN NATURAL SETTINGS

Collecting large samples of diverse respondents in a natural setting is a challenge for our discipline. Although nationally representative surveys can achieve this end, they are generally very expensive. Students, however, often are willing to participate and are far more affordable. Yet, they present at least two concerns for external validity (Mintz, Redd, and Vedlitz 2006; Peterson 2001).

First, student samples are not representative of general adult populations (Oakes 1972; Sears 1986). This concern often is overstated, however, because students tend to resemble adult populations across a range of important covariates, such as partisanship and media use (Druckman and Kam 2011, 51). Moreover, if scholars are interested in estimating relationships between variables, they can use student samples to create valid inferences—even in cases in which the sample differs substantially from the population. If a treatment effect of interest is homogeneous in the population, any

doi:10.1017/S1049096514001127

sample can produce an unbiased estimate. However, even if the treatment effect varies, it can be modeled as long as the sample provides variation across the relevant moderating variables. Thus, unbiased estimates of treatment effects require diverse but not representative samples. For example, in the case of presidential debates, the effect of candidate attention to immigration on viewers' attitudes toward the candidate might depend on a viewer's ideology and race. In this case, unbiased estimates would depend on obtaining a sufficient number of respondents across the ranges of ideology and race but would not require the sample's percentage of conservatives or African Americans (for instance) to equal those in the population (Druckman and Kam 2011). Many single-campus student samples may lack this needed variation.

Second, student-based studies generally are conducted in artificial settings—often a computer lab. Laboratory environments tend to eliminate distractions, resulting in treatment effects that are larger than those in natural settings (Jerit, Barabas, and Clifford 2013). One solution is to allow participation in more natural settings (Kinder 2007) in which distractions introduce variation in participant attentiveness (e.g., Albertson and Lawrence 2009). However, technological and logistical limitations often impede this approach.

Crowdsourcing data collection can mitigate both concerns. A relatively new concept in business and an even newer concept in academia, crowdsourcing is "a strategic model to attract an interested, motivated crowd of individuals capable of providing solutions superior in quality and quantity to those that even traditional forms [can]" (Brabham 2008).[1] Our approach, described in detail below, builds on crowdsourcing work by reaching out to the political science community to access a more diverse student-respondent pool participating in more natural settings. Of added benefit, this approach provides instructors with resources to facilitate classroom discussions—and may even heighten student engagement in the political process.

Figure 1

React Labs: Educate App Interface

(Color online)

COLLEAGUE CROWDSOURCING FOR THE 2012 PRESIDENTIAL DEBATES

Our substantive interest is to understand how candidate debate behaviors affect viewers' attitudes (Boydstun et al. 2014). Despite the salience and visibility of presidential debates (Benoit, Hansen, and Verser 2003; Jamieson and Birdsell 1990; Marcus and Mackuen 1993), few studies have collected real-time reactions that allow for the study of individual debate moments; those that have done so use very small samples (e.g., Fridkin et al. 2007; McKinney and Rill 2009; Pfau, Houston, and Semmler 2005).

Thus, we set out to measure debate reactions using a web application, or "app," that we designed for use on smartphones.[2] The app was also accessible from tablets and personal computers, allowing viewers to react to the debates in real time from anywhere with Internet connectivity. A screenshot of this app, React Labs: Educate, is displayed in figure 1. Respondents used the app while watching the debates live, indicating (at any time they wished) whether they "agreed" or "disagreed" with the candidates and whether they thought the candidates were "spinning" or "dodging" the question.

We needed a larger, more diverse sample of app users than any of our campuses could provide in isolation or combined. Therefore we targeted our recruitment efforts at instructors across the country, knowing that they are uniquely able to encourage student participation (e.g., in exchange for extra credit). To encourage instructors to register their classes and promote participation, we designed an incentive package aimed at helping them to achieve some of their own teaching and learning goals.

Participation far exceeded our expectations, with respondents from all 50 states, the District of Columbia, Puerto Rico, and even outside of the United States.

The materials we provided to registered instructors are available on the project website (http://reactlabseducate.wordpress.com). Before the debates, registered instructors received the following materials:

- PowerPoint slides and lecture notes covering the history of presidential debates—including YouTube links to memorable debate moments as well as research on debate rhetoric, debate strategies, and debate effects
- discussion questions
- a list of resources, websites, and research collections on presidential campaigns and debates
- citations and abstracts of relevant debate research

- alternative assignments for students unable to watch the debates live

After the debates, registered instructors also received the following:

- Within 12 hours of each debate: presentation-ready PowerPoint slides with preliminary results from respondents who used the app
- After the final debate: for each debate, a list of their students who participated

These resources linked political science teaching and research, helping instructors discuss the debates in a way that connected theory with contemporary politics.

We recruited instructors by sending more than 120 individual e-mails inviting colleagues to participate in the project and by sending invitations to key listservs and blogs.[3] Instructors registered their classes to participate through the project website. Each registered course was assigned a unique course identification number, which enabled us to send instructors confirmation of their students' participation but also required us to send a unique e-mail with instructions and the course identification number for each registered class. This challenge was made easier by Gmail's Mail Merge, which allowed us to merge e-mail addresses, course identification numbers, instructor names, and course names from a database into individual e-mails, thereby automating the process of sending individualized messages.[4]

We embedded a predebate survey in the app itself and used a paid (but relatively inexpensive) subscription to SurveyMonkey® to administer a postdebate survey. SurveyMonkey provided the capacity to handle a high volume of student participants, to ask a large number of follow-up questions, and to download the results in a spreadsheet.

Following through on our promise to provide next-day figures and preliminary results proved challenging. We offered our graduate students free food and good cheer to stay up all night after each debate, crunching numbers and compiling PowerPoint slides. Although the process was labor intensive, we felt that providing instructors with immediate results that they could use in class to facilitate discussions of the debates was a critical incentive for participation.

Our research design represents a major advance in external validity. In terms of representativeness, the app allows us to draw on a large and diverse enough sample to include the variation we need for analysis. In terms of artificiality, the app allows students to participate in the study from wherever they would normally watch a debate (e.g., home, a friend's house, or a debate-watch party).

RESULTS

Participation far exceeded our expectations, with respondents from all 50 states, the District of Columbia, Puerto Rico, and even outside of the United States. In total, 263 instructors registered at least one course to participate in at least one debate, representing 361 courses and more than 13,000 potential student respondents.[5] Across the three presidential debates and one vice presidential debate, almost 5,000 undergraduates participated at least once.[6] Counting each respondent in each debate separately, the app received 8,006 respondents, the demographics of which are summarized in table 1.

As table 1 illustrates, our sample is similar to national population means for gender, income, race, party identification, and religion. The major demographic difference is in age because our recruitment efforts were targeted at college undergraduates. Although the sample is not nationally representative, nonetheless we received more than 175 participants in each age group, allowing us to estimate debate effects that vary with age. In terms of both representativeness and variation across a range of variables, these data represent major

Table 1

Study Demographics Compared to National Demographics

	APP		NATIONAL
	N	%	%
Gender[a]			
Women	3,789	48	51
Men	4,099	52	49
Income[a]			
<$25K	1,232	16	18
$25K–$49,999	1,236	16	24
$50K–$74,999	1,397	18	19
$75K–$,99,999	1,140	14	14
>=$100K	2,868	36	26
Race[a]			
African American	694	9	13
Asian	679	9	5
Hispanic	1,054	13	17
Other	418	5	2
White/Caucasian	5,120	64	63
Party ID[b]			
Democratic (includes leaners)	4,215	54	50
Independent	1,235	16	11
Republican (includes leaners)	2,396	31	39
Religion[c]			
Christian	4,737	60	76
Jewish	381	5	1
Muslim	157	2	<1
Atheist or agnostic	2,069	26	15
Other	616	8	8
Age[d]			
18–24	6,830	85	13
25–29	448	6	9
30–39	366	5	17
40–49	183	2	18
>=50	179	2	43

Notes: App estimates include all 8,006 participants across the four debates, including those who participated in more than one debate. The numbers do not total 8,006 on any given demographic item due to non-response on that item.
[a] National estimates are from the US Census.
[b] National estimates are from the Pew Research Center for the People, October 2012, accessed January 23, 2013, from the iPOLL Databank, The Roper Center for Public Opinion Research, University of Connecticut. Available at http://www.ropercenter.uconn.edu/data_access/ipoll/ipoll.html.[7]
[c] National estimates are from the 2008 American Religious Identification Survey.
[d] National estimates are from the 2012 American Community Survey One-Year Estimates.

progress in sample quality over single-campus convenience samples. Table 2 illustrates this variation in more detail.

Part A of table 2 displays the number of students who took part in the debate study, categorized by ideology and race/ethnicity. The table shows that the large number of respondents provided a sufficient number in each cell to model heterogeneous treatment effects—even for those cells that captured rare combinations (e.g., conservative African Americans).

For comparison, part B of table 2 shows the same breakdowns for ideology and race/ ethnicity compiled from the five courses in which students participated from a single campus (University of California, Davis). There are only three African Americans in the UC Davis sample, none of whom identify as conservative, thereby preventing the estimation of heterogeneous treatment effects for this group. This data binning problem occurs across a range of demographic and attitudinal measures.

Thus, our crowdsourcing approach realized several benefits over traditional, single-campus, fixed-location research studies. Although the sample is not representative and app users may have been paying closer attention to the debates than typical viewers, this approach allowed us to collect data in more natural settings than previously possible. It also enables estimates of treatment effects across a range of covariate profiles that otherwise would be inaccessible. Therefore, the sample cannot provide an unbiased estimate of the prevalence of a certain trait in the general population, but it is uniquely suited to produce estimates of many different treatment effects.

THE TEACHING AND LEARNING BENEFITS OF CROWDSOURCING

In addition to the methodological and logistical benefits of our crowdsourcing approach, our solution facilitated teaching and learning. Because of their salience and scale, presidential debates represent key opportunities to encourage student engagement with the political process, which can improve political knowledge and civic skills—especially among those with lower initial levels of political interest (Beaumont et al. 2006). When instructors highlight engagement and civic themes, their students' future political engagement and voter turnout increase (Hillygus 2005; McCartney, Bennion, and Simpson 2013). Furthermore, watching debates tends to boost political efficacy, trust, and information among youth while decreasing cynicism (Kaid, McKinney, and Tedesco 2007; McKinney and Rill 2009). Many of our student participants likely would not have watched the debates were it not for the app and the incentives that we encouraged instructors to offer. Even for those students who would have watched anyway, using our app turned watching TV—a generally passive activity—into an interactive experience. Extensive research has demonstrated that active learning techniques improve test scores (McCarthy and Anderson 2000), engagement with the material (Brown and King 2000; Hess 1999; Ruben 1999; Wolfe and Crooktall 1998), learning (Pace et al. 1990; Perry 1968; Sutro 1985; Washbush and Gosen 2001), and interest (Hess 1999; Smith and Boyer 1996). Although we do not directly measure these effects here, the literature leads us to expect that using the app aided student learning.

We view the teaching benefits of our study—providing instructors with easy-to-use classroom materials and a method by which to actively engage students in the political process—as a hopeful indication that the colleague-crowdsourcing approach can facilitate a symbiotic relationship between teaching and research.

Our crowdsourcing method benefited instructors as well. During the month of October 2012, our publicly available webpage featuring overnight result summaries was accessed more than 5,000 times. In addition to the result summaries, participating instructors accessed our password-protected teaching-resources webpage 450 times. We view the teaching benefits of our study—providing instructors with easy-to-use classroom materials and a method by which to actively engage students in the political process—as a hopeful indication that the colleague-crowdsourcing approach can facilitate a symbiotic relationship between teaching and research.

THE FUTURE OF COLLEAGUE CROWDSOURCING

We believe colleague crowdsourcing holds considerable promise for future studies, particularly in light of ongoing technological innovations, which make national (or even international) crowdsourcing

Table 2

Participant Frequencies by Ideology and Race/Ethnicity

A. ALL APP USERS

	ASIAN	AFRICAN AMERICAN	HISPANIC	CAUCASIAN	OTHER	TOTAL
Liberal	348	368	468	2,080	201	3,465
Moderate	249	262	429	1,390	152	2,482
Conservative	79	60	149	1,606	62	1,956
Total	676	690	1,046	5,076	415	7,903

B. UNIVERSITY OF CALIFORNIA, DAVIS, SAMPLE

	ASIAN	AFRICAN AMERICAN	HISPANIC	CAUCASIAN	OTHER	TOTAL
Liberal	41	2	27	47	15	132
Moderate	27	1	15	25	10	78
Conservative	6	0	2	25	4	37
Total	74	3	44	97	29	247

Notes: Ideology and race were measured in the predebate survey. Ideology was measured with a 100-point sliding scale ranging from 0 (extremely liberal) to 100 (extremely conservative). In the table, participants scoring between 0 and 39 on this scale are classified as liberal, between 40 and 60 as moderate, and between 61 and 100 as conservative.

increasingly feasible. Our app facilitated crowdsourcing by enabling participation across the country, but there are many other potential uses of colleague crowdsourcing; we certainly do not expect all scholars to create an app.

For example, colleague crowdsourcing might be used to foster large-scale and geographically diverse participation in studies using survey platforms such as Qualtrics® and SurveyMonkey®—or participation by specific target groups, such as first-generation college students or Muslims. Colleague crowdsourcing could be used to collect simple cross-sectional survey data, panel data during the course of an academic term, or data derived from survey experiments. It also could be used to measure aspects of the political environment (e.g., counting yard signs or political bumper stickers). Additionally, we can imagine the incentive portion of the crowdsourcing approach taking many forms, including access to the data, webcast guest lecturers, and research notes on the findings for use in class. With enough lead time to include information about a study in their syllabi and/or to incorporate time for discussion in their lecture plans, many instructors may be keen to encourage student participation in an interesting study. In short, the crowdsourcing approach as a recruitment technique is flexible and scalable. Overall, new research technologies coupled with colleague crowdsourcing create a rich opportunity to incorporate research methods, local and global findings, and temporally relevant data in the classroom in a way that can aid research efforts while stimulating a new level of active learning.

ACKNOWLEDGMENTS

We are immensely grateful to the hundreds of instructors and thousands of students who participated in this crowdsourcing exercise; our colleagues Debra Leiter, Jack Reilly, and Michelle Schwarze, who helped develop the teaching resources we used to encourage participation and who sacrificed themselves for our postdebate all-night data-crunching sessions; and our colleague Philip Resnik at the University of Maryland, who conceived of the mobile reactions platform and with whom we collaborated to make React Labs: Educate a reality. We presented a previous version of this article at the 2013 APSA Teaching and Learning Conference. ■

NOTES

1. In the natural sciences, crowdsourcing has yielded considerable payoffs (e.g., the Leafsnap and Tag a Tiny programs). In political science, this model forms the basis for projects such as the Cooperative Congressional Election Study (CCES), the Cooperative Campaign Analysis Project (CCAP), and Time-sharing Experiments for the Social Sciences (TESS).
2. The specific features of the React Labs: Educate app—what it should look like and do—were designed in collaboration with Philip Resnik of the University of Maryland and built using his React Labs technology platform (see Boydstun et al. 2014 for a detailed discussion), with implementation accomplished using a contract development firm. Although the development of mobile apps can be complicated, apps useful for research often can be created at reasonable expense, particularly if one takes a "web app" approach (i.e., apps that run as web pages in device browsers) rather than a "native app" approach (i.e., apps that are programmed for specific devices like iPhones). For researchers with a programming background (or with access to students who have such a background), many websites and software packages make the leap to web app development accessible. For example, http://jquerymobile.com/resources provides an extensive list of resources for jQuery Mobile, one of the most popular mobile client frameworks, and https://docs.djangoproject.com/en/dev/intro is a good starting point for getting up and running with Django, one of the most popular frameworks for implementing the server side in Python. Generally speaking, we suggest contacting a local computer science department as an initial starting point for discussion about app design and availability of programming support. Contract developers also can be found and hired through websites such as oDesk, Freelancer, and Elance. In software development, as for any project, it is important to hire carefully; to set concrete and realistic goals; and to take an incremental, agile approach to the development process.
3. Had Hurricane Katrina not struck, attendees of teaching and learning panels at APSA 2012 would have received lovely color flyers advertising our project; instead, said flyers sit unappreciated in our offices.
4. Several tutorials for Gmail Mail Merge are available online.
5. These 263 instructors included 20 graduate students, 21 non-tenure track faculty, 152 professors, and 70 instructors with some other or nonspecified positions. Using US News designations, their institutions included 121 national universities, 25 national liberal arts colleges, 59 regional universities/colleges, 23 community colleges, 8 international institutions, and 17 other or nonspecified affiliations.
6. Participants include only those respondents who identified their age as 18 or older during the pretest. Participants younger than 18 and those who did not respond to this item are omitted.
7. The survey results reported here were obtained from searches of the iPOLL Databank and other resources provided by the Roper Center for Public Opinion Research, University of Connecticut.

REFERENCES

Albertson, Bethany, and Adria Lawrence. 2009. "After the Credits Roll: The Long-Term Effects of Educational Television on Public Knowledge and Attitudes." *American Politics Research* 37 (2): 275–300.

Beaumont, Elizabeth, Anne Colby, Thomas Ehrlich, and Judith Torney-Purta. 2006. "Promoting Political Competence and Engagement in College Students: An Empirical Study." *Journal of Political Science Education* 2 (3): 249–70.

Benoit, William L., Glenn J. Hansen, and Rebecca M. Verser. 2003. "A Meta-Analysis of the Effects of Viewing U.S. Presidential Debates." *Communication Monographs* 70 (4): 335–50.

Boydstun, Amber E., Rebecca A. Glazier, Matthew T. Pietryka, and Philip Resnik. 2014. "Real-Time Reactions to a 2012 Presidential Debate: A Method for Understanding Which Messages Matter." *Public Opinion Quarterly* 78 (Special Issue).

Brabham, Daren C. 2008. "Crowdsourcing as a Model for Problem Solving: An Introduction and Cases." *Convergence: The International Journal of Research into New Media Technologies* 14 (1): 75–90.

Brown, Scott W., and Frederick B. King. 2000. "Constructivist Pedagogy and How We Learn: Educational Psychology Meets International Studies." *International Studies Perspectives* 1 (2): 245–54.

Druckman, James N., and Cindy D. Kam. 2011. "Students as Experimental Participants: A Defense of the 'Narrow Data Base.'" In *Handbook of Experimental Political Science*, ed. J. N. Druckman, D. P. Green, J. H. Kuklinski, and A. Lupia. New York: Cambridge University Press.

Fridkin, Kim L., Patrick J. Kenney, Sarah Allen Gershon, Karen Shafer, and Gina Serignese Woodall. 2007. "Capturing the Power of a Campaign Event: The 2004 Presidential Debate in Tempe." *The Journal of Politics* 69 (3): 770–85.

Hess, Frederick M. 1999. *Bringing the Social Sciences Alive*. Needham Heights, MA: Allyn & Bacon.

Hillygus, Sunshine D. 2005. "The MISSING LINK: Exploring the Relationship between Higher Education and Political Engagement." *Political Behavior* 27 (1): 25–47.

Jamieson, Kathleen Hall, and David S. Birdsell. 1990. *Presidential Debates: The Challenge of Creating an Informed Electorate*. Oxford: Oxford University Press.

Jerit, Jennifer, Jason Barabas, and Scott Clifford. 2013. "Comparing Contemporaneous Laboratory and Field Experiments on Media Effects." *Public Opinion Quarterly* 77 (1): 256–82.

Kaid, Lynda Lee, Mitchell S. McKinney, and John C. Tedesco. 2007. "Introduction: Political Information Efficacy and Young Voters." *American Behavioral Scientist* 50 (9): 1093–111.

Kinder, Donald R. 2007. "Curmudgeonly Advice." *Journal of Communication* 57 (1): 155–62.

Marcus, George E., and Michael B. Mackuen. 1993. "Anxiety, Enthusiasm, and the Vote: The Emotional Underpinnings of Learning and Involvement during Presidential Campaigns." *The American Political Science Review* 87 (3): 672–85.

McCarthy, J. Patrick, and Liam Anderson. 2000. "Active Learning Techniques Versus Traditional Teaching Styles: Two Experiments from History and Political Science." *Innovative Higher Education* 24 (4): 279–94.

McCartney, Alison Rios Millett, Elizabeth A. Bennion, and Dick Simpson (eds.). 2013. *Teaching Civic Engagement: From Student to Active Citizen*. Washington, DC: American Political Science Association.

McKinney, Mitchell S., and Leslie A. Rill. 2009. "Not Your Parents' Presidential Debates: Examining the Effects of the CNN/YouTube Debates on Young Citizens' Civic Engagement." *Communication Studies* 60 (4): 392–406.

Mintz, Alex, Steven B. Redd, and Arnold Vedlitz. 2006. "Can We Generalize from Student Experiments to the Real World in Political Science, Military Affairs, and International Relations?" *Journal of Conflict Resolution* 50 (5): 757–76.

Oakes, William. 1972. "External Validity and the Use of Real People as Subjects." *American Psychologist* 27 (10): 959.

Pace, David, Bill Bishel, Roger Beck, Peter Holquist, and George Makowski. 1990. "Structure and Spontaneity: Pedagogical Tensions in the Construction of a Simulation of the Cuban Missile Crisis." *The History Teacher* 24 (1): 53–65.

Perry, William G. 1968. *Forms of Intellectual and Ethical Development in the College Years.* New York: Jossey-Bass.

Peterson, Robert A. 2001. "On the Use of College Students in Social Science Research: Insights from a Second-Order Meta-Analysis." *Journal of Consumer Research* 28 (3): 450–61.

Pfau, Michael, J. Brian Houston, and Shane M. Semmler. 2005. "Presidential Election Campaigns and American Democracy: The Relationship between Communication Use and Normative Outcomes." *American Behavioral Scientist* 49 (1): 48–62.

Ruben, Brent D. 1999. "Simulation, Games, and Experience-Based Learning: The Quest for a New Paradigm for Teaching and Learning." *Simulation and Gaming* 30 (4): 498–506.

Sears, David O. 1986. "College Sophomores in the Laboratory: Influences of a Narrow Data Base on Social Psychology's View of Human Nature." *Journal of Personality and Social Psychology* 51 (3): 515–30.

Smith, Elizabeth T., and Mark A. Boyer. 1996. "Designing In-Class Simulations." *PS: Political Science and Politics* 29 (4): 690–4.

Sutro, Edmund. 1985. "Full-Dress Simulations: A Total Learning Experience." *Social Education* 49: 628–34.

Washbush, John, and Jerry Gosen. 2001. "Learning in Total Enterprise Simulations." *Simulation and Gaming* 32: 281–96.

Wolfe, Joseph, and David Crooktall. 1998. "Developing a Scientific Knowledge of Simulation/Gaming." *Simulation and Gaming* 29 (1): 7–20.