

Testing the Effectiveness of a Number-Based Classroom Exercise

Brian Guay, *Duke University, University of Richmond*

Caroline Chandler, *RTI International*

Jennifer Erkulwater, *University of Richmond*

David Landy, *Indiana University, University of Richmond*

ABSTRACT

Large numbers permeate political life, and students of political science can expect to encounter a wide range of numbers in newspaper articles, course readings, and statistics. Recent research in cognitive psychology demonstrates that American adults make systematic errors when comparing numbers in the millions, billions, and trillions. Political decisions made by voters often require weighing large quantities that range across many orders of magnitude, which is difficult without at least a basic understanding of relative magnitudes. If students also lack an understanding of large numbers, professors cannot meaningfully teach them about political phenomena involving such magnitudes. The authors designed and tested an exercise to improve students' accuracy in dealing with large magnitudes, which had immediate and sustained effects on their political judgments about information involving large numbers.

"The proposal will raise 2014 spending for the full fiscal year to \$1.012 trillion. It would save \$85 billion while eliminating \$63 billion in forced spending cuts to the military and other programs through sequestration to achieve total deficit reduction of \$23 billion, budget leaders said."

—*CNN, 2013*

"Lebanon, with a population of less than five million, has taken in more than 1.1 million Syrian refugees, while Jordan has 608,000 and Turkey 815,000....Tens of thousands more fled to Iraq over the past three years only to face new dangers from the onslaught of Islamic militants based in Syria."

—*New York Times, 2012*

American citizens are often called on to make political judgments that require familiarity with large numbers, yet studies find that people make systematic errors in the estimation of large magnitudes. The failure of American citizens to engage with large numbers in a critical manner—especially numbers used

in policy discussions and news media—is a serious impediment to meaningful public discussions about a wide range of policy issues and political matters (Best 2012; Blastland and Dilnot 2007; Huff 1993; Maier and Imazeki 2012). An accurate understanding of large magnitudes not only reinforces better citizenship; it also prepares students for content that is commonly encountered in political science courses, including budget politics, migrant populations, and climate-change statistics. Without a proper facility with these numbers, it is difficult to imagine how students can contribute substantively to discussions about current events or the basic political phenomena that involve interactions across scales (e.g., population change, emissions levels, and the national budget) (Andersen and Harsell 2005; Payne and Williams 2011). Even outside of political science, there has been increased attention given to the effect of numeracy on decision making, personal health choices, and future employment opportunities for students (Peters et al. 2006; Ward, Schneider, and Kiper 2011).

Americans make consistent and predictable errors when dealing with large-number magnitudes, including those in the millions, billions, and trillions—magnitudes that frequently appear in discussions of economic growth, budgeting, and human-population flows (Landy, Silbert, and Goldin 2012, 2013; Resnick et al. 2012). The most commonly used measure of an individual's understanding of number magnitudes is the number-line task, which instructs participants to place numbers on a blank horizontal line with two endpoints (see Link et al. 2013 for a general review). Landy et al. (2013) instructed adult participants to place numbers in the

Brian Guay is a PhD student in American politics at Duke University. He can be reached at brian.guay@duke.edu.

Caroline Chandler is an analyst at RTI International. She can be reached at caroline.e.chandler@gmail.com.

Jennifer Erkulwater is associate professor of political science at the University of Richmond. She can be reached at jerkulwa@richmond.edu.

David Landy is assistant professor of psychology at Indiana University. He can be reached at dlandy@indiana.edu.

thousands and millions range on a number line with endpoints of 1 thousand and 1 billion. They found that 40% of participants consistently and systematically overestimated the magnitudes of these numbers. For example, they placed 1 million nearly halfway between 1 thousand and 1 billion (where 500 million is located). There are 1 thousand millions in a billion, so the correct location of 1 million for this task is immediately to the right of the left endpoint.

Several studies indicate that numeracy can be manipulated and improved, often by brief interventions that improve an individual's ability to estimate large orders of magnitude. Studies with children and adults find that short interventions can improve the ability to estimate orders of magnitude in immediate posttests (Landy et al. 2012; Opfer and Siegler 2007; Resnick et al. 2012; Thompson and Opfer 2010).

Still, surprisingly little literature exists on best practices for teaching number-evaluation strategies to students in the classroom and for integrating those practices across the political science curriculum. A rich pedagogical literature on teaching research methods attests to the trepidation that students experience when encountering numbers in formal ways in political science

other vehicles for stimulating learning (e.g., simulations, discussed later in this article).

CLASSROOM EXERCISE

First, students gained Concrete Experience with numerical magnitudes by physically interacting with large quantities. They were given a 5-foot length of string and asked to hold it tautly in front between their hands. The instructor explained that 1 billion bacteria of an average diameter could fit side by side if lined up on the string. Bacteria were used as the measurable objects because students could physically interact with a representation of not only thousands and millions but also billions of bacteria, which would be difficult with larger objects inside a classroom. Students then were asked to spend several minutes using a colored marker to shade the area of string where 1 million bacteria would fit if lined up side by side. By physically experiencing and interacting with their conceptualizations of large magnitudes and comparing 1 million to 1 billion, they gained a more concrete experience with numbers than a simple conversation might have produced.

By training students to think more accurately about large magnitudes, the exercise also influences their perceptions of political information over an eight-week period.

courses (Bos and Schneider 2009; Sloomaeckers, Kerremans, and Adriaensen 2014; Wagner, Garner, and Kawulich 2011). However, exercises that professors can use to teach orders of magnitude—essential for understanding situations that involve large and small-scale actions—are noticeably missing (Wagner, Garner, and Kawulich 2011). We address this gap by using Kolb's model of experiential learning to design a lesson that will help students grasp and retain numerical information (Kolb 1984).

Drawing from the theories of John Dewey, Kurt Lewin, and Jean Piaget, Kolb's model rejects the view of learning as the accumulation and retention of facts and rather views it as a process in which students ask questions and then deploy the skills needed to acquire the knowledge to answer them (Kolb 1984; Svinicki and Dixon 1987). More specifically, learning is a cycle of four distinct processes—rather than an outcome—that result in comprehensive learning. The cycle of learning starts with Concrete Experience—that is, students' personal experience or concrete involvement in a subject. Experience then stimulates Reflective Observation, in which students examine the meaning of their experience from different perspectives. Through Abstract Conceptualization, students integrate their reflections, making connections among ideas, generalizations, and logical inferences, which then are tested in the process of Active Experimentation. This, in turn, gives rise to experiences that activate anew the learning cycle (Kolb 1984).

Using Kolb's Experiential Learning Theory, we designed an exercise that provides a concrete experience to serve as the starting point for reflection, conceptualized thinking, and experimentation. By training students to think more accurately about large magnitudes, the exercise also influences their perceptions of political information over an eight-week period. The exercise requires less than 10 minutes to complete and can be used in conjunction with

Next, students engaged in Reflective Observation by experiencing the accurate proportion of the string on which 1 million bacteria would fit. The instructor asked them to fold their string in half, explaining that because 1 billion is twice 500 million, 500 million bacteria could fit on half of the string. The instructor then explained that they should have shaded $1/500$ of that half of the string, or $1/1,000$ of the entire string (figure 1). The instructor asked them whether 500 of their shaded regions could fit on the string, which forced those who had overestimated the size of 1 million to visualize the extent of their overestimation. Next, large-number magnitudes were explained: thousands, millions, billions, and trillions expand by multiples of 1 thousand, such that 1 thousand is $1/1,000$ of a million, 1 million is $1/1,000$ of 1 billion, and 1 billion is $1/1,000$ of 1 trillion.

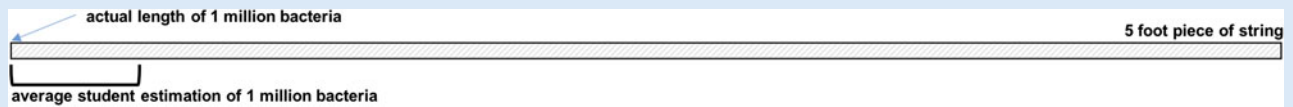
Finally, students interacted with their new knowledge of the relative size of large magnitudes by determining how much string would be covered if 1 million widths of human hair were lined up side by side. At the end of the exercise, the instructor reviewed orders of magnitude by demonstrating how much string 1 thousand hairs would cover and then proceeded to pull an 83-foot length of string into the classroom, explaining that it represented the length of 1 million widths of hair. Students were told that 1 billion pieces of hair would stretch for 16 miles.

TESTING THE EFFECTIVENESS OF THE CLASSROOM EXERCISE

We tested whether the classroom exercise was effective in both increasing students' understanding of large-number magnitudes and changing how they interpret political information involving large numbers—in this case, information involving government spending cuts. We compared their responses to two political judgments and four number-line placements both before and after half of the students were randomly selected to participate in the exercise. We also assessed the long-term effects of the exercise

Figure 1

Perception of 1 Million



Students were told that 1 billion average-sized bacteria could fit side by side on a 5-foot piece of string and were asked to shade the length of string on which 1 million bacteria could fit. This illustration demonstrates the average amount of string (9.74%) shaded by participants compared to the actual length of 1 million bacteria.

with an eight-week retention test, which included a set of political judgments and another number-line task.

A total of 52 undergraduate students in an introduction to American government course at the University of Richmond were invited to participate in the study during class time and were

The intention was to capture cases in which explicit calculations are unlikely; therefore, the endpoint of 1 thousand, rather than zero, was used to discourage them from simply dividing the stimulus number by the right endpoint. Using explicit calculations is one way to comprehend relative magnitudes; however, it is likely that students

We tested whether the classroom exercise was effective in both increasing students' understanding of large-number magnitudes and changing how they interpret political information involving large numbers—in this case, information involving government spending cuts.

compensated \$5 each. Of the 52 students, 31% were in their first year, 46% in their second year, 15% in their third year, and 8% in their fourth year. During the course of the study, the professor was not present in the classroom with the instructors. Students were asked to differentiate between the deficit and the debt, after which all were given the correct answer to create common definitions of terms that appeared later in the study.

Students were given two written excerpts from real political news stories and asked to “rate the impact” of a featured spending cut on an eight-point Likert scale, ranging from “no impact” to “a huge impact”. For example, students were asked to evaluate the impact of a \$58 billion cut to discretionary spending on the \$1.33 trillion federal deficit.¹ The instruction to rate the “impact” of a budget cut was intentionally vague; it allowed them to evaluate the political scenario as they would while reading a newspaper or a textbook, simultaneously judging both the content and size of the cut.² Whereas people may have considered the impact of a budget cut on the public welfare, the instructions specifically asked the students to evaluate the impact of the cut on the budget itself. Although there is no objectively correct answer to whether a 5% budget cut has a “very small impact,” misunderstanding the numbers involved can lead a person to provide a response that is different from one based on accurately estimated numeric quantities.

Students then were given a short number-line task, during which they were asked to place eight numbers (e.g., 280 million) on individual number lines ranging from 1 thousand to 1 billion (figure 2).

and adults use more implicit means to understand magnitudes when they experience numbers in politics, particularly in the media.

Next, students were randomly assigned to either an exercise group or an exposure group based on their student ID number. Those in the exercise group participated in the 10-minute exercise with an instructor. Those in the exposure group, which served as a control group, did not participate in the exercise. After completing the number-line and political-estimation tasks, students in the exposure group watched a short, unrelated video in a separate classroom. It is worth noting that the exposure group was not a simple control because students were still exposed to large numbers during the initial number-line task completed by both groups. They did not receive explicit instruction on magnitudes; however, they may have had enough prior knowledge that simply engaging in the number-line task slightly improved their number-line placements or political judgments involving large numbers.³ Students in both groups then completed an additional number-line task and a set of political judgments immediately after the exercise and in an eight-week posttest.

RESULTS

Number-Line Task

To assess the accuracy of the students' number-line placements, we calculated the mean difference between each placement and its proper location. Doing so assigned a mean overestimation score to each student for each of the three number-line tasks (i.e., pretest, posttest, and eight-week retention).

We examined the effect of the classroom exercise on both number-line placements and political judgments by submitting each to an ANOVA, including time of test (pretest, immediate

Figure 2

The Number-Line Task



The number line task is a cognitive task used to assess number comprehension. Students are instructed to place a number, here 280 million, on a blank number line between 1 thousand and 1 billion.

posttest, and eight-week retention) crossed with exercise condition (exposure or exercise) as independent predictors. Time of test was a within-participants factor, whereas exercise condition was a between-participants factor. We included gender, a five-item political knowledge test (Delli Carpini and Keeter 1996), and political-party identification as covariates. However, we found no effects of these factors on number-line placements or political judgments and do not discuss them further in this article.

The effect of test time ($F(2, 102) = 27.1, p < 0.0001$) on number-line placements was significant in that placements were more accurate

The effect of condition on the political judgments was determined by comparing individual impact ratings across time of test and exercise condition, using the same ANOVA approach described previously: time of test (within participants) and condition (between subjects) were used as independent variables. Students in the exercise group reduced their estimates after the exercise more than those in the exposure group ($F(2, 98) = 3.8, p < 0.05$), whose impact ratings essentially did not change (figure 4). The main effect in the ANOVA was significant ($F(2, 98) = 5.5, p < 0.01$), reflecting the large shift in the exercise group.

Students who experienced the string exercise improved the accuracy of their number-line judgments and shifted their evaluations of political situations at both immediate posttest and eight-week retention.

after both the exercise and the exposure. Students' number-line placements became more accurate from pretest to posttest ($F(1, 49) = 30.1, p < 0.001$) and from pretest to eight-week retention ($F(1, 49) = 42.0, p < 0.0001$) (figure 3). Number placement showed more improvement in the exercise group than in the exposure group, but the difference was not statistically significant. Number-line placement also proved robust over time. The correlation between individual students' placements on the number line at posttest and at the eight-week retention was very high ($r = 0.85, p < 0.001$).

Political Judgments

At pretest, estimates of the impact of cuts were correlated with number-line judgments ($r = 0.30, p < 0.05$), confirming that numerical understanding is involved in political judgments. Indeed, among students who made substantial number-line errors (i.e., greater than 20% of the line), 30% estimated a 5% budget cut as "large" or "huge," whereas less than 10% of the more numerically accurate students did so. This is compatible with previous findings that erroneous number-line judgments are related to larger estimates of numerical impact (Landy, Silbert, and Goldin 2013).

The exposure group lowered its estimate of the political impact of budget cuts at the immediate posttest (simple effects: $F(1, 49) = 10, p < 0.01$), as well as eight weeks later (simple effects, eight-week retention-to-pretest: ($F(1, 49) = 6.5, p < 0.05$)). The change was significant: participants in the exposure group reduced their estimates of political impact by three fourths of a standard deviation (Cohen's $d = 0.75$), which is large even for a targeted classroom exercise. In contrast, the exercise group changed insignificantly (Cohen's $d = 0.02$). In the exposure group, there was no statistically significant difference between responses collected immediately after the intervention and eight weeks later ($F(1, 49) = 0.6, p = 0.45$; $F(1, 49) = 0.4, p = 0.5$).

Summary

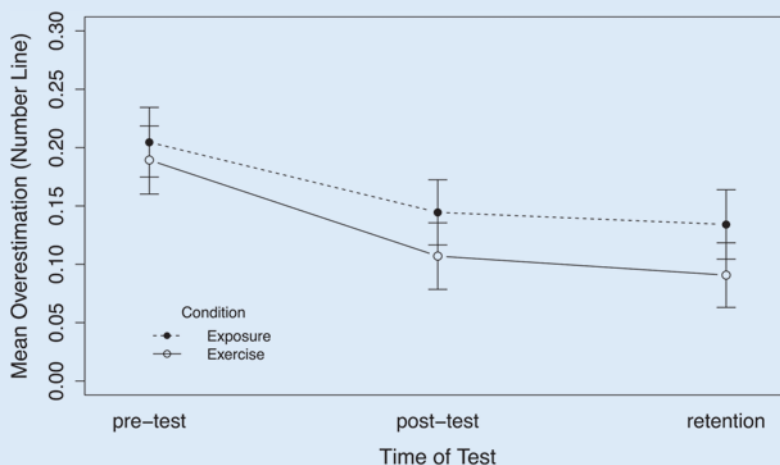
Students who experienced the string exercise improved the accuracy of their number-line judgments and shifted their evaluations of political situations at both immediate posttest and eight-week retention. Mere exposure to large-number estimation (without the exercise) was enough to improve number-line judgments; however, exposure was not sufficient for students to translate that

improved understanding into political contexts. One reasonable interpretation consistent with Kolb's Experiential Learning Theory is that mere exposure to number-line tasks enabled students, after short reflection, to arrive at correct number-line judgments. However, the string exercise helped students to connect this improved understanding with the broader context of political budgets.

GENERAL DISCUSSION AND CONCLUSION

Although we often encounter large numbers in politics, we rarely interact with them in a tangible way. The concrete experience provided by the interactive string exercise is an effective

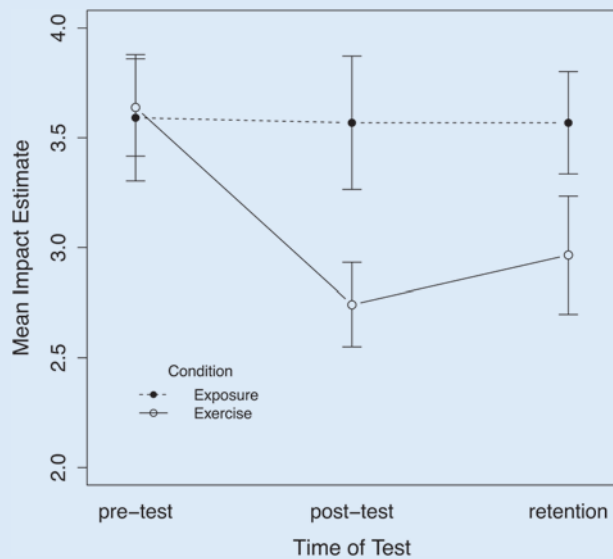
Figure 3
Effect of Exercise on Number Line-Test Accuracy



Both students who received the classroom exercise and who simply gained exposure to numbers by completing the number-line task increased accuracy at both the immediate post-test and eight-week retention.

Figure 4

Effect of Exercise on Political Judgments



Students who received the classroom exercise incorporated new information about large number magnitudes into their interpretation of budget changes. Students who simply received exposure to the number line task did not significantly change how they evaluated these situations.

and worthwhile teaching tool. We found that simply exposing students to the number-line task improved performance, which suggests that actively interacting with large magnitudes can improve their understanding of large numbers. In the process of shading and folding the string, students test their knowledge and are given a physical representation of their own misconceptions of numbers. They then apply their improved understanding of magnitudes to the political-judgment task. When educators use these numbers in the classroom without explicitly dealing with the magnitudes involved, they risk overlooking a significant learning opportunity for their students to concretely experience these quantities, learn to use them accurately, and reconfigure their understanding of the relationships between large magnitudes in a way that is transportable to politically relevant contexts.

The short classroom exercise described in this article can be extended and adapted to provide critical and useful information about large-number magnitudes in other lessons in political science. When teaching about fiscal policy and the budget, the exercise can precede a budget simulation, such as the Concord Coalition's "Principles and Priorities" exercise, in which students are asked to cut various programs or raise taxes to balance the budget (Concord Coalition 2015). During lessons on climate-change policy, the exercise can involve students in experiencing the population of nations and projected population growth in Organization for Economic Co-operation and Development and developing nations, followed by a simulation, such as the Carbon Mitigation Initiative's "Stabilization Wedge Game," in which students must agree to a climate-change accord (Carbon Mitigation Initiative 2015). The exercise described here can complement and strengthen the impact of student simulations and other lessons by effectively providing a more accurate understanding of large-number magnitudes frequently used in political science.

ACKNOWLEDGMENTS

The authors thank Ernest McGowen and Dan Palazzolo for their support and Arthur Charlesworth, Zach Davis, Megan DeLaunay, and Erin Ottmar for their useful feedback. ■

NOTES

1. All spending cuts were intentionally less than 5% of the total budget or deficit being reduced. Although there are no "small" budget cuts, these cuts are smaller than some large-scale reforms proposed by fiscal conservatives.
2. Because all students saw the same questions and because we used a within-subjects design, differences between how students perceived the term "impact" are minimally consequential for our current investigation. If an individual student's definition of the impact rating scale remains consistent throughout the study, differences between how students perceive items on the scale (e.g., "no impact" and "large impact") also are minimally consequential.
3. During previous face-to-face pilot studies, participants completing the same number-line task frequently expressed that they learned through the process of placing numbers on the number line. Participants who incorrectly placed 720 thousand on the righthand side of the number line, for example, often realized their overestimation error when asked to place 800 million on the number line.

REFERENCES

- Andersen, Kristi, and Dana Michael Harsell. 2005. "Assessing the Impact of Quantitative Skills Course for Undergraduates." *Journal of Political Science Education* 1 (1): 7–27.
- Best, Joel. 2012. *Damned Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists*. Los Angeles: University of California Press.
- Blastland, Michael, and Andrew Dilnot. *The Tiger That Isn't: Seeing Through a World of Numbers*. London: Profile.
- Bos, Angela L., and Monica C. Schneider. 2009. "Stepping around the Brick Wall: Overcoming Student Obstacles in Methods Courses." *PS: Political Science & Politics* 42 (2): 375–83.
- Carbon Mitigation Initiative. 2015. "Stabilization Wedges." Princeton, NJ: Princeton University. Available at <http://cmi.princeton.edu/wedges>. Accessed October 13, 2014.
- Concord Coalition. 2015. "Principles and Priorities." Washington, DC: Concord Coalition. Available at www.concordcoalition.org/act/tools/principles-priorities. Accessed September 20, 2014.
- Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven, CT: Yale University Press.
- Huff, Darrell. 1993. *How to Lie with Statistics*. New York: W. W. Norton.
- Kolb, David A. 1984. *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ: Prentice Hall.
- Landy, David, Noah Silbert, and Aleah Goldin. 2012. "Getting Off at the End of the Line: The Estimation of Large Numbers." Proceedings of the 34th Annual Conference of the Cognitive Science Society. Sapporo, Japan. August, 1–4.
- . 2013. "Estimating Large Numbers." *Cognitive Science* 37: 775–99.
- Link, Tanja, Stefan Huber, Hans-Christoph Nuerk, and Korbinian Moeller. 2013. "Unbounding the Mental Number Line: New Evidence on Children's Spatial Representation of Numbers." *Frontiers in Psychology* 4: 1021.
- Maier, Mark H., and Jennifer Imazeki. 2012. *The Data Game: Controversies in Social Science Statistics*. Armonk, NY: M. E. Sharpe.
- Opfer, John E., and Robert S. Siegler. 2007. "Representational Change and Children's Numerical Estimation." *Cognitive Psychology* 55: 169–95.
- Payne, Geoff, and Malcolm Williams. 2011. *Teaching Quantitative Methods: Getting the Basics Right*. London: Sage Publications, Inc.
- Peters, Ellen, Daniel Västfjäll, Paul Slovic, C. K. Mertz, Ketti Mazzocco, and Stephan Dickert. 2006. "Numeracy and Decision Making." *Psychological Science* 17 (5): 407–13.
- Resnick, Ilyse, Thomas F. Shipley, Nora Newcombe, Christine Massey, and Theodore Wills. 2012. "Examining the Representation and Understanding of Large Magnitudes Using the Hierarchical Alignment Model of Analogical Reasoning." Proceedings of the 34th Annual Conference of the Cognitive Science Society. Sapporo, Japan. August 1–4.
- Slootmaeckers, Koen, Bart Kerremans, and Johan Adriaensens. 2014. "Too Afraid to Learn: Attitudes Towards Statistics as a Barrier to Learning Statistics and to Acquiring Quantitative Skills." *Politics* 34 (2): 191–200.

Svinicki, Marilla D., and Nancy M. Dixon. 1987. "The Kolb Model Modified for Classroom Activities." *College Teaching* 35 (4): 141–6.

Thompson, Clarissa A., and John E. Opfer. 2010. "How 15 Hundred Is Like 15 Cherries: Effect of Progressive Alignment on Representational Changes in Numerical Cognition." *Child Development* 81 (6): 1768–86.

Wagner, Claire, Mark W. J. Garner, and Barbara Kawulich. 2011. "The State of the Art of Teaching Methods in the Social Sciences: Towards a Pedagogical Culture." *Studies in Higher Education* 36 (1): 75–88.

Ward, R. M., M. C. Schneider, and J. D. Kiper. 2011. "Development of an Assessment of Quantitative Literacy for Miami University." *Numeracy* 4 (2): 1–19.

APPENDIX

Political Impact Questions

"The nonpartisan commission advised that through executive action, congressional rule, and legislation, a number of steps be taken immediately to show that Washington can lead by example. These collected recommendations would reduce spending on both the security and nonsecurity sides of the discretionary budget. Together, they will save more than \$50 billion in 2012 alone." [The projected deficit for 2012 is \$1.33 trillion.] What was the impact of the spending reduction on the projected 2012 deficit?

"Congress has emphasized that nearly every corner of the government is on the chopping block, and its newly passed proposal includes a reduction of \$194 million in spending for the legislative branch." [The budget for the legislative branch in 2011 was \$5.54 billion.] What was the impact of this proposal on the legislative branch's 2011 budget?

"Republican lawmakers took the hatchet to funding for their own legislative branch with a cut of \$103 million." [The budget for the legislative branch in 2011 was \$5.54 billion.] What was the impact of these cuts on the legislative branch's 2011 budget?

"Congressmen and Congresswomen on the House Appropriations Committee took what its chair called a 'wire brush' to the budget, proposing the removal of \$58 billion in nonsecurity discretionary budget funds." [The federal budget deficit in 2011 was \$1.33 trillion.] What was the impact of the removal of nonsecurity discretionary funds on the 2011 federal deficit?