Testing the Effectiveness of a Number-based Classroom Exercise

Brian Guay[1,2]

Caroline Chandler[1]

Jennifer Erkulwater[1]

David Landy[1,3]

[1]University of Richmond; [2]Duke University; [3]Indiana University

Article accepted for publication in *PS: Political Science & Politics*

Keywords: classroom exercises, quantitative skills, political behavior

Abstract

Large numbers permeate political life; 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 has demonstrated 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 students about political phenomena involving such magnitudes. Therefore, we designed and tested an exercise to improve students' accuracy in dealing with large magnitudes, which had immediate and sustained effects both on their abilities to accurately work with large numbers and their perceptions of numbers in political judgments.

"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 (Desjardin & Walsh, 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 (Benhabib, 2012)

American citizens are often called upon 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; Maier and Imazeki 2012; Huff 1993). An accurate understanding of large magnitudes not only reinforces better citizenship, but also prepares students for content that is commonly encountered in political science courses, such as budget politics, migrant populations, and climate change statistics. Without a proper facility with these numbers, it is hard to imagine students could contribute substantively to discussions of current events or the basic political phenomena that involve interactions across scales, such as population change, emissions levels, or the national budget (Payne and Williams 2011; Andersen and Harsell 2005). Even outside of political science, there has been increased attention paid to the effect of numeracy on decision-making, personal health choices, and future employment opportunities for students (Peters et al. 2006; Ward et al. 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, or human population flows (Landy et al. 2012, Landy et al. 2013, Resnick et al. 2012, Guay et al. under review). 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 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, placing 1 million, for example, nearly halfway between 1 thousand and 1 billion (where 500 million is located). Since there are 1 thousand millions in a billion, the correct location of 1 million on this task is just to the right of the left endpoint.

Several studies indicate that numeracy can be manipulated and improved, often by very 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 (Thompson & Opfer 2010; Opfer & Siegler 2007; Landy et al. 2012; Resnick 2012).

Still, surprisingly little literature exists on best practices for teaching number evaluation strategies to students in the classroom, or for integrating these practices across the political science curriculum. A rich pedagogical literature on teaching research methods attests to the trepidation that students feel about encountering numbers in formal ways in political science courses (Bos and Schneider 2009; Wagner et al. 2011; Slootmaeckers et al. 2014). However, exercises that professors can use to teach orders of magnitude—essential for understanding situations that involve large-scale and small-scale actions—are noticeably missing (Wagner et al. 2011). We address this gap by using David Kolb's model of experiential learning to design a lesson that will help students grasp and retain numerical information.

Drawing from the theories of John Dewey, Kurt Lewin, and Jean Piaget, David Kolb's model rejects the view of learning as the accumulation and retention of facts, but rather as a process in which students ask questions and then deploy the skills needed to acquire the knowledge to answer those questions (Kolb 1984; Svinicki and Dixon 1987). More specifically, learning is not an outcome but a cycle of four distinct processes that result in comprehensive learning. The cycle of learning starts with Concrete Experience, a student's personal experience or concrete involvement in a subject. Experience then stimulates Reflective Observation, in which the student examines the meaning of his or her experience from different perspectives. Through Abstract Conceptualization, the student integrates his or her reflections, making connections between ideas, generalizations, and logical inferences, which are then 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 about large magnitudes more accurately, the exercise also influences students' perceptions of political information over a substantial time period (8 weeks). The exercise took less than 10 minutes to complete and can be used in concert with other vehicles for stimulating learning, such as simulations, which we discuss later.

Classroom Exercise

First, students gained Concrete Experience with numerical magnitudes by physically interacting with large quantities. Students were provided a five-foot length of string and asked to hold it in front of them taut 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 measureable objects because students were able to physically interact with a representation of not only thousands and millions of bacteria, but also billions, which would be difficult to do with larger objects inside a classroom. Students were then given a colored marker and asked to spend several minutes shading the area 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, students gained a more concrete experience with numbers than a simple conversation about the difference between 1 million and 1 billion might have produced.

Next, students engaged in Reflective Observation by experiencing the accurate proportion of the string that 1 million bacteria would fit on. The instructor asked students to fold their string in half and explained that because 1 billion is twice 500 million, 500 million bacteria could fit on half of the string. The instructor then explained that students should have shaded 1/500 of that half of the string, or 1/1000 of the entire string (see Figure 1). The instructor asked students 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 to students; thousands, millions, billions, and trillions expand by the power of 1 thousand, such that 1 thousand is 1/1000 of a million, 1 million is 1/1000 of 1 billion, and 1 billion is 1/1000 of 1 trillion.

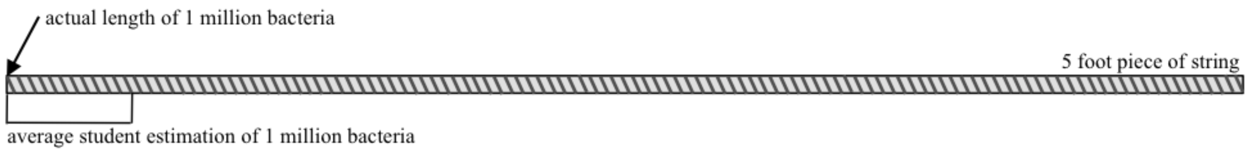

Figure 1. Students were told that 1 billion average sized bacteria could fit side by side on a five foot piece of string, and were asked to shade the length of string that 1 million bacteria could fit on. The illustration above demonstrates the average amount of string, 9.74%, shaded by participants compared to the actual length of 1 million bacteria.

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 very end of the exercise, the instructor reviewed orders of magnitude by demonstrating how much string 1 thousand hairs would cover and 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 learned 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 students interpret political information involving large numbers, in this case information involving government spending cuts. We compared students' responses to two political judgments and four number line placements both before and after half of students were randomly selected to participate in the exercise. We also assessed the long-term effects of the exercise with an eight-week retention test, which included set of political judgments and another number line task.

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

Students were given two written excerpts from real political news media and asked to 'rate the impact' of a featured spending cut on an 8-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 non-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 students to evaluate the political scenario as they would while reading a newspaper or textbook, simultaneously judging both the

[1] All spending cuts were intentionally less than 5% of the total budget or deficit being reduced. While there is no such thing as a "small" budget cut, these cuts are smaller than some large-scale reforms proposed by fiscal conservatives.

content and size of the cut[2]. While people may have considered the impact of a budget cut on the public welfare, the instructions specifically asked them to evaluate the impact of the cut on the budget itself. While there is no objectively right answer to whether a 5% budget cut has a 'very small impact', misunderstanding the numbers involved can lead a person to a provide different response than if they accurately estimated numerical quantities. Previous evidence (and evidence from this experiment, reported below) shows that people who overestimate more on number lines also make larger political impact estimates (Landy, Silbert, & Goldin, 2013; Guay et al., under review).

Students were then 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 (see Figure 2). The endpoint 1 thousand, rather than zero, was used to discourage students from simply dividing the stimulus number by the right endpoint, since the intention was to capture cases in which explicit calculations are unlikely. While explicit calculations are one way to comprehend relative magnitudes, it is far more likely that students and adults use more implicit means of understanding magnitudes when they experience numbers in politics, particularly in the media.

---

[2] Since 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 the 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", "large impact") are also minimally consequentially.

Figure 2. The number line task is a cognitive task used to assess number comprehension. After receiving instructions, students here are instructed to place 280 million on a blank number line between 1 thousand and 1 billion.

Next, students were randomly assigned to either an *exercise group* or *exposure group* based on their student ID numbers. Students in the *exercise group* participated in the 10 minute exercise with an instructor. Students 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* is not a simple control because students were still exposed to large numbers during the initial number line task completed by both groups. Although they did not receive explicit instruction on magnitudes, they may have had enough prior knowledge of magnitude systems that a small prompt would slightly improve their number line placements or political judgments involving large numbers[3]. Students in both the *exercise group* and the

---

[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. A participant who incorrectly places 720 thousand on the right-hand side of the number line, for example, often realized their overestimation error when asked to place 800 *million* on the number line.

exposure group then completed an additional number line task and set of political judgments immediately afterwards and at an eight-week post-test.

Results

Number line task

To assess the accuracy of students' number line placements, we calculated the mean of the difference between each placement and its proper location. This assigned a mean overestimation score to each student for each of the three number line tasks (pre, post, and 8-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 (pre-test, immediate post-test, and eight-week retention) crossed with exercise condition (exposure, exercise) as independent predictors. Time of test was a within-participants factor, while exercise condition was between-participants. We included gender, Delli Carpini and Keeter's five-item test (Delli Carpini and Keeter 1996), and political party identification as covariates, though we found no effects of these factors on number line placements or political judgments and will not discuss them further.

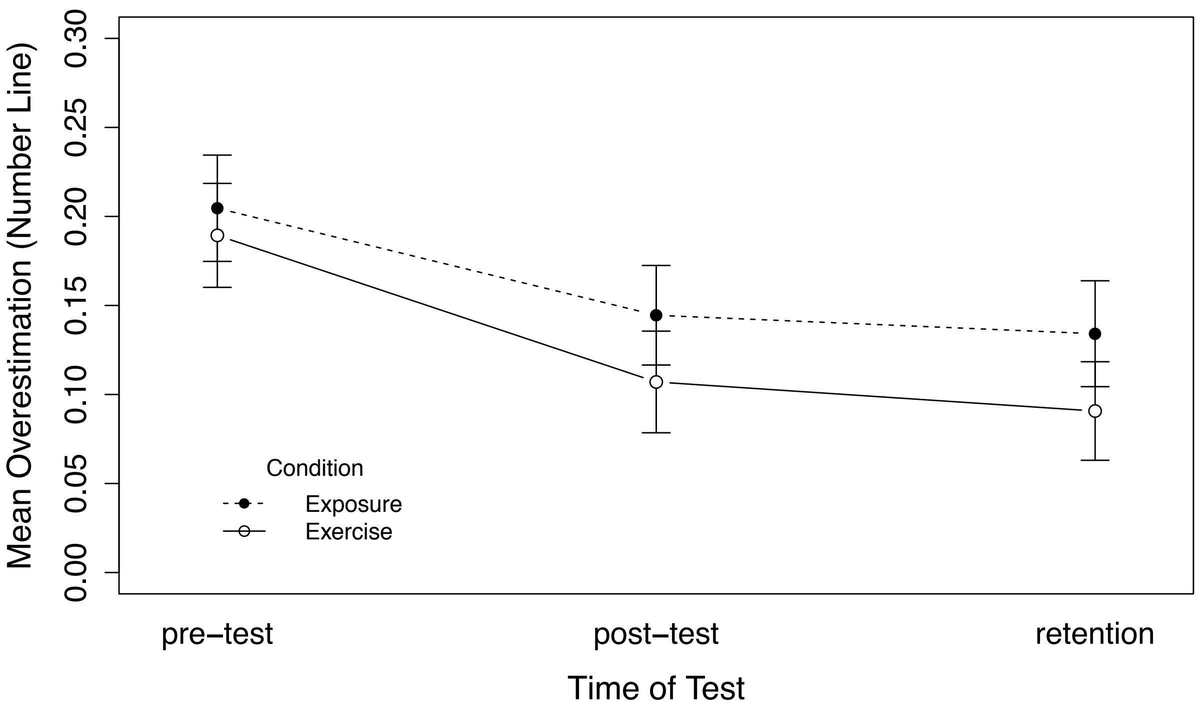

Figure 3: Both students who received the classroom exercise and who simply gained exposure to numbers by completing the number line task increased accuracy at both the immediate post-test and eight-week retention.

The effect of test time ($F(2, 98) = 27.1$, $p<0.0001$) on number line placements was significant, such that placements were more accurate after both the exercise and exposure. Students' number line placements became more accurate from pre-test to post-test ($F(1,49)=30.1$, $p<0.001$), and from pre-test to retention ($F(1,49)=42.0$, $p<0.0001$, see Figure 3). Number placement improved numerically more 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 student's placements on the number line at post-test and at the retention was very high ($r=0.85$, $p<0.001$).

Political Judgments

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

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

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, retention-pre: ($F(1,49)=6.5$, $p<0.05$)). 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). There was no significant change in the political judgments of the exposure group. The change was quite large: 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 quite large even for a targeted classroom exercise. In contrast, the exposure group changed very little (d=0.02), and not significantly.

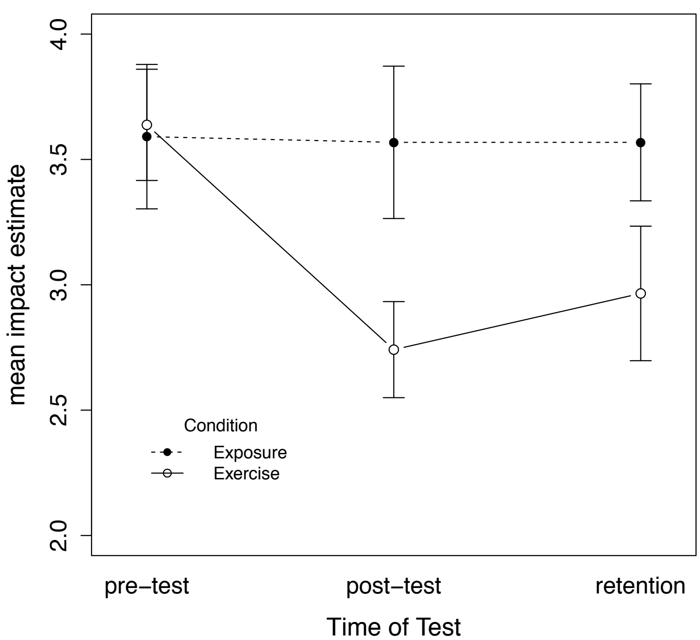

Figure 4: Students who received the classroom exercise incorporated their new information about large number magnitudes into their interpretation of budget changes. Students who simply received exposure to the number line task did not significantly change how they evaluated these situations.

Summary

Students who experienced the string exercise both improved in their number line judgments and shifted their evaluation of political situations, both at immediate post-test and 8-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 theory of experiential learning is that mere exposure to number line tasks enabled people to, after short reflection, come to correct number line judgments, but that the string exercise served to help students connect this improved understanding with the broader context of political budgets.

General Discussion and Conclusion

Though we encounter large numbers often in politics, we rarely interact with them in a tangible way. The concrete experience provided by the interactive string exercise serves as an effective and worthwhile teaching tool. We found that simply exposing students to the number line task improved performance on that task, suggesting that actively interacting with large magnitudes can improve students' 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. Students then apply their improved understanding of magnitudes to the political judgment task. When educators use these numbers in the classroom without dealing explicitly 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 understandings of the relationships between these large magnitudes in a way that is transportable to politically relevant contexts.

The short classroom exercise presented here can be extended and adapted to provide students with critical and useful information about large number magnitudes in other lessons in political science. While teaching the budget and fiscal policy, the exercise can precede a budget simulation, such as the Concord Coalition's *Principles and Priorities*, in which students are asked to cut various programs or raise taxes in order to balance the budget (Concord Coalition 2014). During lessons on climate change policy, the exercise can involve students experiencing the population of nations and projected population growth in OECD 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). This exercise compliments and strengthens the impact of student simulations and other lessons by effectively providing students with a more accurate understanding of number magnitudes used frequently in political science.

Appendix:

Political Impact Questions

The non-partisan commission advised that through executive action, congressional rule, and legislation, a number of steps be taken immediately to show Washington can lead by example. These collected recommendations would reduce spending on both the security and non-security 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 their newly passed proposal includes a reduction of $194 million in spending for the legislative branch. [The legislative branch's budget 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 legislative branch's budget in 2011 was $5.54 billion] What was the impact of these cuts on the legislative branch's 2011 budget?

Congressmen on the House Appropriations Committee took what its chairman called a 'wire brush' to the budget, proposing the removal of $58 billion in non-security discretionary budget funds. [The federal budget deficit in 2011 was $1.33 trillion] What was the impact of the removal of the non-security discretionary funds on the 2011 federal deficit?

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.

Benhabib, Seyla. "The Morality of Migration." 2012. New York Times, July 29. http://opinionator.blogs.nytimes.com/2012/07/29/stone-immigration/.

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 and Politics 42 (2): 375–83.

Carbon Mitigation Initiative. 2015. "Stabilization Wedges." Princeton, NJ: Princeton University, on-line document at http://cmi.princeton.edu/wedges/.

Concord Coalition. 2015. "Principles and Priorities." Washington, DC: Concord Coalition, on-line document at http://www.concordcoalition.org/act/tools/principles-priorities.

Delli Carpini, Michael X., and Scott Keeter. 1996. What Americans Know about Politics and Why It Matters. New Haven, CT: Yale University Press.

Desjardins, Lisa and Deirdre Walsh. 2013. “Budget Deal Aims to Avert Another Shutdown.” CNN, December 10. http://www.cnn.com/2013/12/10/politics/budget-negotiations/.

Guay, Brian M., Zach J. Davis, Megan DeLaunay, Art Charlesworth, and David Landy. Under Review. “Number Comprehension Impacts Political Judgments”.

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: Prentice Hall.

Landy, D., 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.

Landy, David., Noah Silbert and Aleah Goldin. 2013. “Estimating large numbers.” Cognitive Science 37: 775–799.

Maier, Mark H. and Jennifer Imazeki. 2012. The Data Game: Controversies in Social Science Statistics. Armonk: M.E Sharpe.

McKeough, Anne, Judy Lee Lupart, and Anthony Marini. Teaching for transfer: Fostering generalization in learning. Routledge, 2013.

Opfer, John E., Robert S. Siegler. 2007. “Representational change and children’s numerical estimation.” Cognitive Psychology (55): 169-195.

Payne, Geoff and Malcolm Williams. 2011. Teaching Quantitative Methods: Getting the Basics Right. London: Sage.

Peters, Ellen, Daniel Viastfjall, Paul Slovic, C.K. Mertz, Ketti Mazzocco, and Stephan Dickert. 2006. “Numeracy and decision making”. *Psychological Science* 17 (5):407−413.

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.

Slootmaeckers, Koen, Bart Kerremans, and Johan Adriaensen. 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): 141-146.

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): 1768-1786.

Wagner, Claire, Garner, Mark W.J., and Kawulich, Barbara. 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., Schneider, M. C., & Kiper, J. D. 2011. “Development of an assessment of quantitative literacy for Miami University”. Numeracy, 4(2): 1-19.