Journal of Political Science Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/upse20

Promoting Productive Political Dialogue in Online Discussion Forums

Aaron S. King, J. Benjamin Taylor & Brian M. Webb

To cite this article: Aaron S. King, J. Benjamin Taylor & Brian M. Webb (2021) Promoting Productive Political Dialogue in Online Discussion Forums, Journal of Political Science Education, 17:sup1, 724-750, DOI: 10.1080/15512169.2020.1808476

To link to this article: https://doi.org/10.1080/15512169.2020.1808476

JOURNAL OF POLITICAL SCIENCE EDUCATION
2021, VOL. 17, NO. S1, 724–750
https://doi.org/10.1080/15512169.2020.1808476

Routledge
Taylor & Francis Group

SCHOLARSHIP OF TEACHING AND LEARNING

Promoting Productive Political Dialogue in Online Discussion Forums

Aaron S. King[a], J. Benjamin Taylor[b], and Brian M. Webb[c]

[a]University of North Carolina Wilmington; [b]Kennesaw State University; [c]Gordon State College

ABSTRACT

Discussion is a crucial component for learning in a college classroom. Increasingly, university and college faculty are using online learning management systems to facilitate and assess course discussions. Given this reality, are there ways to frame prompts to generate normatively better discussions, or discussions where students are better able to meet the course learning outcomes? To answer these questions, we utilize data from Introduction to American Government classes at two institutions with the students of three instructors who participated in online discussion boards on multiple substantive topics; for each topic, students were randomly assigned to one of two experimental conditions. In each, we framed a related prompt in different ways to test how such prompts impact student success as measured by several learning outcomes informed by the American Association of Colleges & Universities (AAC&U). With our unique experimental design and novel data, we are able to test several hypotheses related to student engagement, the content of discussions, as well as the quality of students' work. In the end, our research has important normative implications for pedagogy as well as the cultivation of civility and political engagement inside and outside of the modern classroom environment.

ARTICLE HISTORY
Received 25 October 2019
Accepted 3 August 2020

KEYWORDS
Civic education; asynchronous discussions; online teaching; mixed-method research

The university model is predicated on a learned professor leading her students through a guided discussion of complex topics and questions where the students not only interact with the professor but their peers as well. Since Plato, there is a tradition of discourse in Western-style education. This is particularly the case in higher education where high-minded strategies like the "Socratic Method" and critical thinking are the hallmarks of a college education. In the modern context, discussion takes place in areas beyond the physical classroom. Online discussion boards are a valuable pedagogical tool and demonstrate the capacity to make the classroom experience richer when utilized in face-to-face (F2F) courses (e.g., Bliuc et al. 2010; Dabbagh and Bannan-Ritland 2005; Krentler and Willis-Flurry 2005; Rovai 2007; Williams and Lahman 2011).

Discussions are particularly valuable when thinking about political issues, which are often value-laden. Successful students can use political science as a lens through which

CONTACT J. Benjamin Taylor jtayl369@kennesaw.edu Political Science, Kennesaw State University, 1000 Chastian Road, Kennesaw, GA 30144-5588, USA.
Authorship is equal and listed alphabetically.

Supplemental data for this article can be accessed at www.tandfonline.com/upse

to engage with political phenomena. With the analytical and empirical approach espoused by the scientific method, we believe these discussion forums give students an opportunity to improve their critical thinking skills by applying their knowledge to important topics within the discipline. Online discussions provide a vehicle to improve students' skills and understanding of the substantive content of the course, but they can also help form the foundation for healthy civic dialogue after students leave campus.

Given the state of technology in higher education and our view on what makes a student successful, what strategies can faculty use to generate meaningful, content-rich discussions via learning management systems (LMS) in college classes?[1] Moreover, are there specific effects when considering issues discussed in introductory political science classes? To answer our research questions, we use prompt framing for online discussion boards to assess several dependent variables using students in our Introduction to American Government courses.

As part of a traditional F2F course, eight low-stakes discussion forums engaged students with course content on a range of topics.[2] Student-subjects were randomly assigned to either a control condition with "traditional" framing language, or a treatment condition that includes "priming" language.[3] For our purposes, our priming language is intended to bring to front-of-mind something specific to a topic that is likely to generate affective engagement by our students. Our design allows us to test three sets of dependent variables related to (1) student engagement, (2) the content of discussions, and (3) student success. Our use of an experimental design to assess the effects of discussion boards offers insight for political science education, but it also has implications for teaching and learning research more broadly.

First, we briefly outline the extant knowledge on the value of online discussions and the role of framing, then we motivate three learning outcomes of interest. Next, we present our theoretical expectations and specific hypotheses. We then describe our unique experimental design and discuss our data. In the end, we find mixed results concerning student engagement as a function of discussion topic and experimental condition. We do find significant evidence that the content of boards varies across topics, and priming language increases the negative sentiment of discussion posts. After taking differences between instructors into account, we find some instances where certain topics and priming language can impact student success as measured by rubric scores; in some cases, a treatment condition decreased scores by 7 percentage points and, in another case, it increased the scores by 8 percentage points. Finally, we close by considering the implications of our research for students and their instructors.

Online discussion as a pedagogical tool

Online discussion forums are an increasingly used tool in modern American higher education (Jones and Jones 2014) due to the ease with which forums can be populated, the ways they can give students a more dynamic way to interact outside the classroom, and an increased attention on "writing-to-think" exercises (Backer 2016; Hamann, Pollock, and Wilson 2012; Wilson, Pollock, and Hamann 2007). Despite increasing use, scholarship notes practitioners' ambivalence; either instructors deride forums as an unhelpful hindrance or praise them for enlivening and expanding content for students

in both F2F and online environments (Brookfield and Preskill 2012; Morris and Stommel 2013). Like any pedagogical tool, discussion forums delivered via LMS are only as good as the instructor's designs for their implementation, which vary with the level and nature of the course (Janssens-Bevernage 2014).

A typical method to generate discussion is the "post-once, reply-twice" rule. While some suggest this method is not the best way to generate an organic give-and-take from students online for more mundane topics (see Janssens-Bevernage 2014), it is a good way to ensure students actively participate. Importantly, prompts for the forums themselves need to be considered when one assesses the "liveliness" of the discussion. Students often fall into one-word responses or generally lack enthusiasm when these minimum requirements are the guide for their participation, which negatively affects both their experience (Wilson, Pollock, and Hamann 2014) and their professor's experience (McGuire 2016). An important incentive to get genuine, engaged discussion in an online environment is the framing of the discussion forum itself.

Framing is a widely studied and fundamental consideration for anything related to discussion, political or otherwise (Druckman 2011). Frames affect the political choices people make even when those choices are empirically identical (Kahneman and Tversky 2000). Framing by politicians and other elites can affect politics in myriad ways such as in citizens' understanding of foreign policy issues (Entman 2004), domestic policy issues (Bartels 2005, 2007; Bolsen 2013), and the way the public interprets news media and other information (e.g., Dilliplane 2011; Levendusky 2013; Taylor 2017). Here, we focus on the potential impact of frames in the classroom when they are embedded within online discussions.

Frames are crucial for discussion posts because the set the bounds for the assignment, which inform the students about the limits of acceptable discourse. Courses and workshops on best practices in pedagogy make it clear that for students to perform in ways that meet the expectations of faculty, it is incumbent on faculty to frame their assignments in a way that makes the topic evident and thought provoking (Davis 2017). This is particularly true for online discussions where the framing choices instructors make affect the enjoyment students express as well as their overall performance (Wu and Hiltz 2004). Thus, we use different frames as the key experimental intervention to assess how students perform in online discussion forums.

Learning outcomes for online discussion forums

In this project, discussion forums are the pedagogical vehicle, but the learning outcomes we assess are (1) the degree of student engagement, (2) the content of the discussion, and (3) the quality of the discussion, or student success as defined by our rubric. These dependent variables are important because they represent concepts of interest for both teaching and learning research as well as political science more generally, but of course there are other learning outcomes instructors, academic departments, and institutions may value.[4] As instructors, it is important to design courses and assignments with the learning outcomes in mind, and to reflect on the types of assignments that generate the qualitatively best student achievement.

Student engagement and the content of political discussions

As opposed to passive learning activities like listening to a lecture or reading a textbook, online discussions offer students the opportunity for active learning by engaging with the material and interacting with their peers in a structured environment (Bonwell and Eison 1991; Hamann, Pollock, and Wilson 2009). Yet, in online environments that are frequently asynchronous, it is more challenging to create the "community" that is present in the typical F2F classroom (Glazier 2016). Students must take greater initiative to engage with the assignment, which may be driven by the substantive topic of the discussion or the prompt itself.

Our empirical interest is rooted in seeing how discussion prompts can improve student success, and success is more likely with increased levels of student engagement. While engagement is important, the content of such engagement—the content of a discussion, or discussion type—hinges on the sentiments expressed within each post. To assess this, we ascertain whether traditional or priming prompts generate different levels of engagement as well as different types of affective responses. Measuring sentiment for online discussion boards is quite common for businesses (Homburg, Ehm, and Artz 2015), and we know that general political discussion is heavy with affective components (Marcus 2002; Marcus, Neuman, and MacKuen 2000). In political science, measuring sentiment has been a great tool to better understand political communication (Young and Soroka 2012), especially through social media platforms. For example, research shows that Twitter behavior can serve as a useful proxy for ideology (Barberá 2015; King, Orlando, and Sparks 2016).

Measuring sentiment in discussion boards is important because political theorists place a premium on civility in political discussion. Civility is a norm to be exemplified in the classroom not just because it is pleasant, but because it represents the highest ideal in a democratic society: everyone's positions are respected, heard, understood, and judged fairly (e.g., Barber 1999, 2003). Furthermore, we know framing is a crucial mechanism for generating differences in response types across a variety of political topics and in diverse political settings (e.g., Bartels 2005; Druckman 2011). This project extends this research into the political science classroom.

Discussion quality and student success

The quality of political discussion—both in society and in the classroom—is a constant source of angst for those who study political communication. To achieve the highest form of deliberative ideals, there must be an open forum where discussants understand the shared facts of the discussion, engage in a way that allows everyone to make their case, and note where and why there is agreement and disagreement at the conclusion (Fishkin 1991; Mutz 2006). In F2F deliberative environments, social pressure, proximity, and temporal constraints affect willingness to engage (Mutz 2006). This is a less present problem for online forums where people are more willing to engage because the medium gives the needed social distance and time to reflect over the course of the discussion (Stanley and Weare 2004).

"Success" can be measured in a variety of ways. The primary place to assess student success is on the direct outcomes for the assignment (i.e., student grades).[5] Rubrics are

one of the more successful assessment instruments instructors in higher education can use because they give clarity to the students as well as instructors for the relevant assignment outcomes (Reddy and Andrade 2010; Stevens and Levi 2013). Importantly, rubrics give students insight into the nuance of their achievement assuming the rubric is successfully framed and clear about succeeding levels of performance (Reddy and Andrade 2010). Later in our analysis, we use rubric scores as a tool to measure and convey student success in online discussions.

Student success can also be measured as something less related to course outcomes and more akin to best practices for higher levels of achievement.[6] One best practice a student may engage is posting more frequently in discussions. While evidence is mixed on the extent to which frequency, *per se*, generates higher outcomes, we do know that students who engage less often are less likely to be successful in courses where these assignments exist (Davies and Graff 2005). Another best practice is engaging with the discussion earlier in the availability timeframe. Research demonstrates that the initial post in a discussion forum is crucial for generating cognitively complex discussion from respondents (Hara, Bonk, and Angeli 2000), so it may also be the case that those who post earlier are similarly crucial to generating quality discussion engagement. Student success, however defined or measured, is a paramount concern for any instructor, regardless of the course or the field. Yet, in our case, we have a particular interest in understanding how framing choices in online discussion boards affect introductory political science classes. Not only does this allow us to explore measures of student success, but also how alternative frames create differences in sentiment and discussion about politics in the classroom.

Theoretical expectations and hypotheses

Given the preceding literature, our theoretical expectation is that different topics as well as discussion post framing will impact student engagement, the content of the discussion, and the quality of discussion. This will happen because, given the affective nature of topics covered in Introduction to American Government classes, priming or leading frames will produce engagement and greater student success in ways that more traditional or even-handed frames will not.[7] As students engage in various discussion topics and different frames for these topics, we test the following broad hypotheses:

Discussion topic hypotheses

Hypothesis 1: Different topics lead to different levels of *engagement* in online discussion forums.

Hypothesis 2: Different topics lead to different *content* in online discussion forums.

Hypothesis 3: Different topics lead to different *quality* of content in online discussion forums.

Discussion frame hypotheses

Hypothesis 4: Topics that include priming language will lead to greater *engagement* compared to topics framed in a traditional manner.

Hypothesis 5: Topics that include priming language will result in different *content* than topics framed in a traditional manner.

Hypothesis 6: Topics that include priming language will encourage higher *quality* responses compared to topics framed in a more traditional manner.

In a survey course, it is not surprising that students will enjoy some topics more than others, and their work in online discussions will be a function of that interest. Despite having a grade tied to their work, we expect the general topic (e.g., Congress, the Presidency, etc.) and the specific topic (e.g., congressional approval, presidential expectations, etc.) of each discussion will elicit different levels of engagement. Similarly, the content and its quality will also vary. Specifically, we are curious about the substantive content of each discussion (i.e., verbiage and major themes) as well as the sentiment of the posts. We are also interested in the quality of the contributions. We operationalize quality discussions by considering the success of each student, which we measure and convey using a rubric. Conceptually, we believe quality contributions occur when students introduce unique ideas and arguments which they support with material from the class as well as outside evidence.

Importantly, we do not make directional predictions in our first three hypotheses. While topics matter for engagement, content, and quality, we remain agnostic to the direction of these expectations. For example, while we typically expect students will be more engaged discussing campaigns and elections compared to thinking about the Constitution, current events and the salience of issues ebb and flow. The former is more interesting during an election year while the latter topic might be more interesting when there is a salient debate on guns or states' rights. Because our project focuses on pooled data from students in multiple settings, any specific directional expectations would vary across time and space; examining specific expectations conditional on the discussion context is certainly an area that future research should consider.

Although one goal of introductory political science courses is that students will be able to engage with one another with an open mind to consider ideas and positions different from their own, decades of research show us that framing has important consequences in political science. We know that politicians can frame issues to their advantage, the framing of survey questions can influence responses, and the media can frame current events to cater to its intended audience (e.g., Druckman 2011; Kahneman and Tversky 2000). Similarly, we believe the framing of online discussion topics can have important consequences as students craft their responses with the framed prompt in mind. Specifically, we expect the priming frame will lead to greater engagement; however, we recognize this engagement may manifest itself in different ways.

Conceivably, the primes we include may activate students' personal beliefs, leading them to participate as partisans or ideologues in much the same way some individuals participate in the "Comment Section" of online newspaper articles (Coe, Kenski, and Rains 2014; Muddiman and Stroud 2017); research also suggests the social setting in which individuals apply their partisanship matters as well (Klar 2014). To some students, the priming language we use in online discussion prompts may be perceived as more provocative than the language students read in the control prompt. As such, some students may be less likely to form a coherent argument or cite evidence from course material or outside sources to buttress their claims.

On the other hand, we believe priming certain details related to the prompt will benefit students who have a solid understanding of the paradigm of political science and empiricism, which a political science classroom encourages. Rather than distracting them from the task, these priming frames may lead to greater engagement. In this controlled academic setting, a debate about normative issues can take place while still stressing the skills students learn as political scientists. A more politically charged debate may encourage students to form more cohesive arguments and locate convincing evidence to support their claims compared to students who are tasked with a more mundane discussion prompt. Conversely, more partisan or ideologically charged topics could cause students to react more instinctively without regard for evidence or support; this is an empirical question we can answer with this analysis.

In short, we expect student engagement and success will improve as a result of discussion frames that include priming language. In the following section, we discuss our methodological approach for testing our expectations regarding student engagement, discussion board content, and the quality of students' contributions and provide a summary of our data.

Data and methods

To test our hypotheses, we utilize eight 2×2 post-test experimental designs for discussion forums within Introduction to American Government courses at two institutions in the southeastern US in the fall of 2017 and spring of 2018. For each topic, students were randomly assigned to a condition by our LMS before the discussion forums opened.[8]

As shown in the first column of Table 1, the prompts focused on the following eight topics: Campaigns & Elections, Civil Rights, Congress, Constitution, Media, Political Parties, Presidency, and Public Opinion. The second column shows the control condition, or traditional frame, and the third column displays the treatment condition, or the priming frame. Note the text in boldface represents the priming language of the treatment condition. For example, consider the Presidency board which addresses the relationship between presidential expectations and performance. Here, the treatment condition primes students: "Thinking about the Trump presidency … " While students in the control group may think of President Trump when they read the prompt, we believe priming a specific office holder, notably one who is particularly polarizing, results in a discussion prompt that is more politically charged. While political scientists often think about the office of the presidency as an institution, as suggested in the traditional frame, priming Trump may activate strong feelings which will result in students thinking more about the individual who holds the office.[9] To be sure, these framing differences are subtle, so to the extent we see significant relationships in the analysis to follow, they speak to the important consequences frames can have even in a controlled, academic setting.[10]

Table 2 shows a summary of our sample, which includes students of three different instructors over two semesters at two different universities (seven total class sections, including one restricted to Honors students). In sum, 239 students participated in this study leading to 4,486 discussion posts, which serves as our primary unit of analysis.[11,12]

Table 1. Discussion board prompts, by topic and experimental condition.

Topic	Control (traditional frame)	Treatment (priming frame)
Campaigns and elections	In the news, you have heard about the controversy regarding Voter ID Laws. Considering the arguments on both sides, do you think voters should be required to present a photo ID to cast a vote?	In the news, you have heard about the controversy regarding Voter ID Laws. **Some argue there is too much potential for fraud while others claim Republicans favor such laws to prevent some citizens that typically vote for Democrats from exercising their right to vote, particularly minorities and low income voters.** Do you think voters should be required to present a photo ID to cast a vote?
Civil rights	States vary in their laws regarding felons and the right to vote. Under what circumstances do convicted felons criminals lose civil rights? What arguments are made on both sides, and which do you find more persuasive?	States vary in their laws regarding felons and the right to vote. Under what circumstances do convicted felons criminals lose civil rights? **Is it fair that these laws disproportionately impact minorities and lower income citizens?**
Congress	Over 90% of House incumbents win reelection while Congress has an approval rating of less than 20%. How is this possible? What can we do to improve the quality of representation we receive from Congress?	Over 90% of House incumbents win reelection while Congress has an approval rating of less than 20%. **Is this a sign that Congress is broken? Are career politicians so entrenched that the public has no choice but to reelect them even if they aren't happy with their performance?** What can we do to improve the quality of representation we receive from Congress?
Constitution	What would the founders think about the country today, especially in its relationship to the Constitution which they composed?	**Thinking about the current partisan political climate and the Trump presidency**, what would the founders think about the country today, especially in its relationship to the Constitution which they composed?
Media	How do you think the increasing number of outlets for political news affects American politics? Should Americans seek out information from traditional sources? What are the consequences of this?	How do you think the increasing number of outlets for political news affects American politics? Should Americans seek out information from traditional sources? What are the consequences of this? **Are partisan news sources like Fox News and MSNBC bad for our democracy?**
Parties	According to Duverger's Law, it is difficult for third parties to emerge in countries with institutional arrangements like the United States. Do you believe the two political parties present voters with an acceptable level of choice in elections?	According to Duverger's Law, it is difficult for third parties to emerge in countries with institutional arrangements like the United States. **Given the level of partisanship and negativity in our current environment**, do you believe the **Republicans** and **Democrats** present voters with an acceptable level of choice and competition in elections?
Presidency	As Americans, we demand so much from our presidents that they have a hard time living up to our expectations. Is there anything presidents can do to keep the American public satisfied?	As Americans, we demand so much from our presidents that they have a hard time living up to our expectations. **Thinking about the Trump presidency**, is there anything presidents can do to keep the American public satisfied?
Public opinion	Some people believe public opinion, politicians, and our politics are very polarized—particularly among our politicians and political elites. Consider both sides of the argument on polarization and discuss: How polarized are the public, and who is to blame?	Some people believe public opinion, politicians, and our politics are very polarized—particularly among our politicians and political elites. **Some have even gone so far to say we are in the midst of a culture war for the soul of America, particularly on social and racial issues.** Consider both sides of the argument on polarization and discuss: How polarized are the public and who is to blame **(media, interest groups and ideological extremists, big money, etc.)? Thinking about the political climate from your perspective, are we in a "culture war"?**

The composition of our student sample enhances the generalizability of our results. Our sample includes significant gender diversity (51.6% female), and students from all class levels (44.8% freshmen, 39.5% sophomores, 13.0% juniors, and 2.7% seniors) and a variety of academic majors (just 3.1% are political science majors). With this diverse student sample, we can help satisfy external validity concerns such as treatment effects being specific to one setting, and this allows for differences between LMS environments

Table 2. Summary of discussion board participants.

Instructor	Semester	Institution	Section	# of students	Total posts, by section	Posts per student	Posts per student, by instructor
A	Fall 2017	X	Honors	19	460	24.2	26.3
A	Spring 2018	X	1	22	620	28.2	
B	Fall 2017	X	1	48	1025	21.4	21.4
C	Fall 2017	Y	1	33	548	16.6	15.9
C	Fall 2017	Y	2	35	655	18.7	
C	Spring 2018	Y	1	41	650	15.9	
C	Spring 2018	Y	2	41	528	12.9	

as well.[13] Our primary concern is the effect of framing, and the internal validity of our experimental design is enhanced by randomization.

Importantly, these discussion boards—the content and themes—are fully integrated within each course. Though these are online discussions, we use evidence from political science and current events as well as in-class activities and readings to lead into these assignments. These discussions serve as an opportunity for students to apply what they have learned in the course to important topics that are relevant to political scientists and to a democratic society more broadly. Each board was open for one week which allowed students to participate over several days while incorporating class material and their own experiences into their posts.[14]

To test the hypotheses stated above, we proceed using several methods and strategies. While we take an initial look at differences across topics, our primary concern is the role of framing and the consequences stemming from using traditional prompts versus prompts which include priming language. To gain insight into the level of *engagement* as well as the *content* of our discussion prompts and frames, we utilize descriptive statistics, word clouds, and sentiment analysis. Then, we assess the *quality* of their content and the level of student success using rubric scores as the dependent variables in a series of regression analyses.

Analysis and results

Analysis of discussion board engagement

As an initial test of our engagement hypotheses—that engagement varies by topic (H1) and treatment condition (H4)—we present descriptive statistics of our discussion board posts in Table 3, which summarizes just under 4,500 discussion board posts from 239 students split across 8 different topics. The table is split into three sets of columns, including "Total Posts per Student," "Average Words per Post," and "Average Sentences per Post." For each measure, there is a column for all students and each experimental condition. Finally, the "Effect" column indicates a positive or negative treatment effect ("Treatment"—"Control"). Across topics, there is little difference in the "Total Posts per Student"; on average across all topics, each student submitted 3.1 posts.[15] The Presidency topic resulted in the most posts per student (3.24) while the Civil Rights topic resulted in the fewest posts per student (3.01). The next two sets of columns in

Table 3. Descriptive statistics of engagement in discussion boards, by topic.

	Total posts per student				Average words per post				Average sentences per post			
	Total	Control	Treatment	Effect	Total	Control	Treatment	Effect	Total	Control	Treatment	Effect
Campaigns and elections	3.2	3.1	3.3	0.1	161.3	161.8	160.7	−1.1	6.7	6.7	6.6	−0.1
Civil rights	3.0	3.0	3.0	0.0	161.7	177.5	146.6	−30.9	6.8	7.4	6.2	−1.2
Congress	3.2	3.3	3.1	−0.2	161.2	157.3	165.3	8.0	6.9	6.8	7.1	0.4
Constitution	3.1	3.2	3.0	−0.2	168.0	162.9	173.7	10.8	6.8	6.4	7.2	0.8
Media	3.0	3.1	3.0	−0.1	172.5	160.0	185.6	25.6	7.2	6.9	7.4	0.5
Political parties	3.2	3.2	3.2	0.0	162.6	167.4	157.4	−10.0	6.6	6.9	6.3	−0.6
Presidency	3.2	3.1	3.4	0.2	160.1	164.9	155.8	−9.1	6.8	7.1	6.4	−0.7
Public opinion	3.1	3.2	2.9	−0.3	170.1	163.3	178.0	14.6	7.5	7.2	7.8	0.7
All topics	3.1	3.1	3.1	0.0	164.7	163.8	165.6	1.8	6.9	6.8	6.9	0.1
Initial posts					295.9	300.4	291.3	−9.1	12.1	12.3	11.8	−0.5
Replies					103.0	102.4	103.6	1.2	4.5	4.4	4.5	0.0

Table 3 address the length of the posts, including the "Average Words per Post" and the "Average Sentences per Post." As noted in the bottom two rows of the table, students were much more engaged in "Initial Posts" as opposed to "Replies" to their classmates (295.9 words and 12.1 sentences vs. 103 words and 4.5 sentences).[16]

Looking across experimental conditions in Table 3, there is virtually no difference in the number of "Total Posts per Student" (approximately 3.1 in both experimental conditions "Control" and "Treatment" columns), however, there are some treatment effects depending on the substantive topic of the discussion. We predicted students in the treatment group who respond to the priming frame would show greater engagement (H4); the evidence is mixed given only half the topics follow our expectation. The strongest positive treatment effect is the Media topic for "Average Words per Post" (25.6 more words in the treatment condition) and Constitution for "Average Sentences per Post" (0.8 more sentences in the treatment condition). Conversely, the strongest negative effect for both measures is the Civil Rights board (30.9 fewer words and 1.2 fewer sentences in the treatment condition). In the next section, we examine the content of the discussion boards, which provides an additional opportunity to analyze student engagement.

Analysis of discussion board content

Although there are only small differences in the length of the posts depending on the substantive topic or the experimental condition, we now investigate how these online communities vary in content (H2 for topics and H5 for experimental conditions). Word clouds are one tool to efficiently visualize the prominence, or frequency, of words within our discussion boards, and they provide us with an initial snapshot of their content. After cleaning the text of all 4,486 posts from our complete dataset, we use the *wordcloud* (Fellows 2018) package in R.[17]Figure 1 shows the word cloud for all posts within the Media topic, which asked students to speak on the potential consequences of different outlets. The most common word used in these posts was "news," which appeared 1,822 times. For comparison, the term "bias" was used 377 times; as a result, its presence in the word cloud is less pronounced.

While we can compare word clouds across topics, we can also compare the content of the experimental conditions. Consider the Presidency topic where students are primed to think about the Trump presidency in the experimental condition. We present these word clouds in Figure 2, and while they look comparable, there are important differences. For example, it is not surprising that the term "Trump" appears on the treatment word cloud; students mentioned "Trump" 267 times in the treatment condition, but only 61 times in the control condition (for comparison, "Obama" was used 25 in the control and 35 times in the treatment group).

The *wordcloud* package can also create comparison and commonality clouds. As shown in the top panel in Figure 3, this comparison cloud for the Constitution topic compares the control and treatment group within the same cloud with the size of words proportional to the maximum deviation between the frequency of a word in one group compared to its average use in both groups.[18] The comparison cloud presents a much clearer picture of the experimental effect of the prompt. For example, the treatment

Figure 1. Media word cloud.

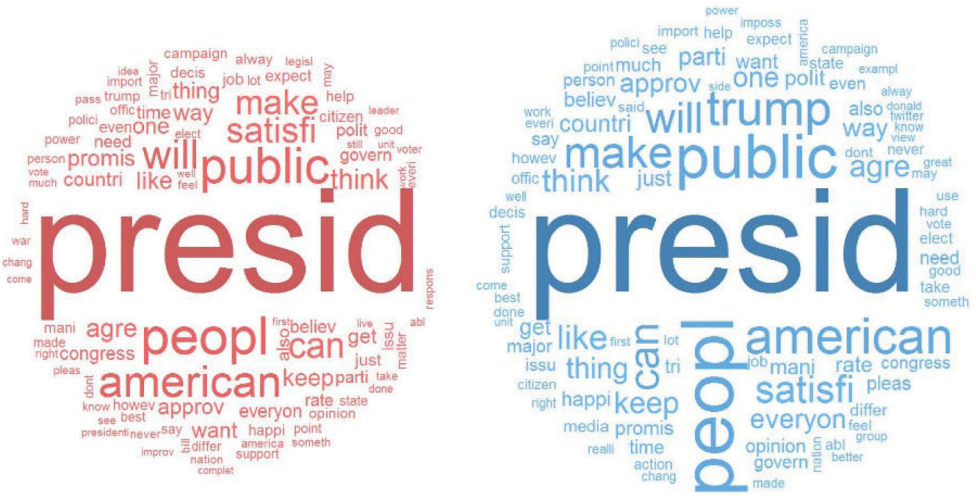

Figure 2. Presidency word cloud, by experimental condition. *Note*: From left to right, this figure presents word clouds for the control and treatment condition.

discussion board focuses much more on Trump and partisan politics as well as specific issues (i.e., immigration) compared to the control discussion board.[19] The bottom panel of Figure 3 shows the commonality cloud, which displays the words used in both the control and treatment groups sized according to their frequency.[20]

While we present several word clouds here, it seems clear that the content varies across topics as well as experimental conditions (H2 and H5).[21] In addition, students do use the words and phrases prompts present to them, which we should not take for

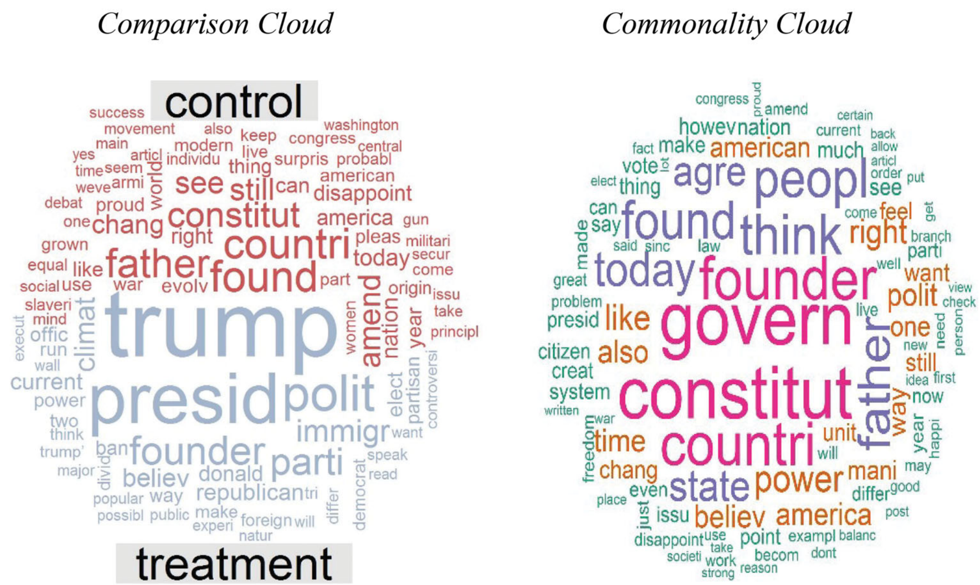

Figure 3. Constitution word clouds.

granted. If students are participating as partisans or ideologues as opposed to students in a political science class, it is conceivable they may be more likely to stray from the prompt presented to them. This is also important because to the extent instructors want their students to become familiar and comfortable with the language of political science, framing prompts using that vocabulary is a critical first step for fostering the meaningful, content-rich discussions we desire.

Using a set of tools collectively known as sentiment analysis, we further investigate the content of these topics (H2) and frames (H5).[22] To understand the sentiment and emotions within the discussions, we use a particular lexicon from The National Research Council (NRC) Word-Emotion Association Lexicon (EmoLex), which includes over 14,000 words classified by eight emotions (i.e., anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) as well as positive and negative sentiment (Mohammad and Turney 2013).[23] Affective engagement is a well-known component for learning in both educational (e.g., Immordino-Yang and Damasio 2007) and political contexts (e.g., Marcus, Neuman, and MacKuen 2000). While political scientists strive for objectivity in their analyses, there is little doubt that when undergrads discuss politics these emotions and sentiments can emerge in spite of the academic setting.

Prior to analyzing our posts, we first transform our dataset from one with each post serving as a unit of analysis to one where every word is a data point; our full dataset with 4,486 posts becomes one with 672,902 words. After running our data through the NRC Lexicon, 110,458 words were categorized into at least one of the emotions or sentiments (16.4%).[24] Just over one-third of these words suggest a "positive" sentiment ($n = 42{,}074$ or 38.1%) while around one quarter suggest a "negative" sentiment ($n = 25{,}472$ or 23.1%). As an example of one of the emotions in the NRC lexicon, 13.3% ($n = 14{,}719$) of the words classified in our discussion boards evoke "anger."

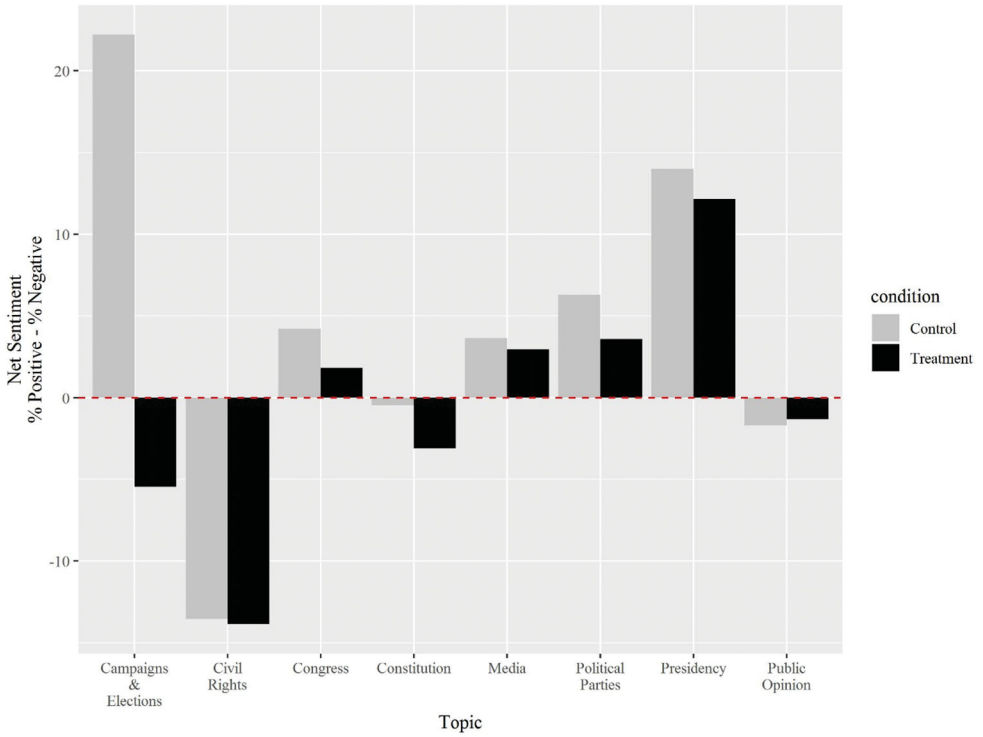

Figure 4. Net sentiment, by topics and experimental condition.

Looking across topics, some discussion boards had a higher percentage of their content classified into these different sentiments. As a percentage of the total words, the Civil Rights boards had the largest share classified (35.4%) while the Media boards had the smallest share classified (18.9%). To draw comparisons between the sentiments of different topics and experimental groups, we construct a measure of "Net Sentiment," which is the percent of positive words minus the percent of negative words, which we show in Figure 4.[25] We draw several conclusions from this figure. First, the Net Sentiment varies greatly across different topics. While a majority of topics are more positive than negative, the Civil Rights and Constitution boards are exceptions. Second, there are differences in the Net Sentiment between the control and treatment groups. In all discussion board topics, the treatment—where students respond to a frame that includes priming language that is excluded from the control group's prompt—leads to a Net Sentiment that is more negative.[26] Interestingly, the control group for the Campaigns & Elections board is quite positive (61.1% positive versus 38.9% negative) whereas the students in the treatment group have a negative Net Sentiment (52.7% negative versus 47.3% positive).

One possible response to these findings is that our priming frames could be harming students if they are reacting more negatively to the prompts. We do not believe this is the case. Rather, as indicated in much of the affective intelligence literature (Immordino-Yang and Damasio 2007; Marcus et al. 2005; e.g., see Marcus, Neuman, and MacKuen 2000) learning can be enhanced by mild negative emotionality.

Analysis of discussion board quality and student success

To help our students better understand how their work was evaluated, to ensure consistency and fairness in the grading process within each class, and in an effort to have our three instructors grade the discussion boards in a consistent manner, we created a rubric for our discussion board assignments based on the VALUE Rubrics from the Association of American Colleges & Universities.[27] The rubric includes six learning outcomes of interest (see Table A1 in the Supplemental Appendix for the complete rubric): Ideas, Arguments & Analysis (IAA), Connection to Course Materials (CCM), Outside Evidence and Support (OES), Contributions to the Learning Community (CLC), Writing Quality (WQ), and Required Postings/Timeliness (RPT). Ranging from Exemplary (100%) to Unsatisfactory (60%), the instructor evaluates the work of each student for every discussion board topic (i.e., eight times throughout the semester), which includes their initial post as well as replies to classmates. In the analysis to follow, we use the six learning outcomes, in addition to the students' final grade as dependent variables. We present descriptive statistics for these variables in the top panel of Table 4 for all categories. The top row presents the "Average Score" for each rubric category and the overall "Grade" for the discussion board assignments, which is 85.20%. Students perform the best in the IAA category (87.28%) and the worst in the RPT category (82.30%).

We hypothesized that different topics and experimental conditions lead to different content quality (H3 and H6). To assess whether the frame that includes priming language improves student success, we use the rubric scores and overall grade as the dependent variables in a series of OLS regressions presented in the bottom panel of Table 4.[28] First, we look for systematic differences based on the course instructor (Instructor A, B, or C) and discussion board topic. Because students should review their grades and improve their performance and the instructors taught the material in a difference sequence, we control for the number of the discussion assignment ("Discussion Order," 1 through 8) in addition to using an indicator variable for the first discussion board assignment within each section.[29]

Compared to Instructor C (baseline instructor), the first two instructors usually gave students higher scores across rubric categories. Depending on the rubric category and topic, there are some grade differences. For example, the scores for the Presidency topic are statistically significantly lower in 4 of the 7 models. The rubric category where the topic is most influential to the final grade is "Outside Evidence and Support." Compared to the baseline topic (Campaigns and Elections), scores for Congress, Media, and Presidency are lower ($p \leq .01$). As a substantive example, consider the OES scores for the Presidency. Students score 6.45 percentage points lower than the students in the same rubric category in the Campaigns and Election discussion. In terms of the controls for the timing of the assignment, the grades for the first discussion boards are no different from the remaining seven topics. The order of the discussions matters only for the OES category; students earn an additional point in this category for each consecutive assignment. Pedagogically, this suggests students are including additional outside evidence as the semester progresses, which is one of the reasons to be transparent in the grading process.

Next, we consider how the inclusion of priming language in a prompt impacts rubric scores. We first add an indicator variable to the model specification above for students

Table 4. Rubric scores for content quality.

Descriptive statistics	IAA	CCM	OES	CLC	WQ	RPT	Grade
Average score	87.28	86.15	86.19	83.35	85.97	82.30	85.20
Standard deviation	10.61	10.69	12.67	11.82	9.73	11.72	9.55
Minimum score	60	60	60	60	60	60	60
Maximum score	100	100	100	100	100	100	100
OLS regression	**IAA**	**CCM**	**OES**	**CLC**	**WQ**	**RPT**	**Grade**
Instructor A	5.42***	0.59	1.40*	4.50***	2.60***	2.00**	2.76***
Instructor B	5.07***	2.50***	−5.65***	5.02***	3.19***	3.83***	2.33***
Instructor C				Baseline instructor			
Campaigns and elections				Baseline topic			
Civil rights	0.39	−0.20	0.42	1.04	1.91*	0.20	0.46
Congress	−2.29*	−1.12	−4.67***	−0.88	−0.72	0.66	−1.60
Constitution	−1.67	0.00	−1.47	0.10	−0.19	−0.46	−0.75
Media	−0.68	0.05	−4.38***	−0.91	0.62	−1.19	−1.20
Political parties	−0.04	1.38	−1.24	0.66	1.30	−0.16	0.18
Presidency	−2.27*	−2.02*	−6.45***	−0.16	−0.58	−0.10	−2.14**
Public opinion	1.01	0.13	−1.97	−1.06	1.41	−1.18	−0.47
Discussion order	0.21	0.50	0.99***	0.16	0.16	0.07	0.25
First assignment	2.32	0.32	−0.11	0.78	0.84	−0.72	0.47
Constant	84.59***	85.48***	85.15***	80.73***	83.51***	81.16***	83.66***
R^2	0.06	0.02	0.07	0.04	0.03	0.02	0.02

Note: Two tail hypothesis tests where: *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Rubric Category Abbreviations: IAA: Ideas, Arguments, and Analysis; CCM: Connection to Course Material; OES: Outside Evidence and Support; CLC: Contribution to the Learning Community; WQ: Writing Quality; RPT: Required Postings/Timeliness.

Table 5. Priming frames and rubric scores for student success, by instructor.

	Instructor		
Rubric category (DV)	A	B	C
IAA		0.77*	
CCM	1.85**		
OES			−1.21*
CLC		1.36**	
WQ	0.96*		
RPT	1.95*		
Grade	1.02*		

Note: Each entry summarizes a different model where the rubric category (DV) is regressed on the treatment condition as well as controls for the topic of the discussion board (not shown). We calculate each of the seven models for each instructor. Note the table only shows entries for significant coefficients (one tail tests) where: $^{*}p < 0.10$; $^{**}p < 0.05$.

Rubric Category Abbreviations: IAA: Ideas, Arguments, and Analysis; CCM: Connection to Course Material; OES: Outside Evidence and Support; CLC: Contribution to the Learning Community; WQ: Writing Quality; RPT: Required Postings/Timeliness.

assigned to the treatment condition. When all students and instructors are combined, the coefficient on the treatment variable is never significant. Because rubric scores consistently vary depending on the instructor, we run these models separately for each instructor, as summarized in Table 5.[30] These models test for a direct effect of the treatment condition while controlling for the discussion board topic (controls not shown). Note that Table 5 includes only the significant coefficients for the Treatment variable, as well as indicators of the level of significance (one-tail tests).

There are four significant treatment effects for Instructor A and two significant treatment effects for Instructor B, all in the expected direction. For example, in Instructor A's class, and holding the substantive topic of the discussion constant, students in the treatment condition receive an additional 1.95 points compared to students in the control condition for the "Required Posting & Timeliness" category (RPT). There is one significant treatment effect for Instructor C, but it is in the opposite direction as our expectation. Here, students in the treatment condition received 1.21 fewer points than students in the control condition in the OES category, which assesses the use of outside sources.

Next, we examine each rubric category with similar models in Table 6, but here we run models separately by the substantive topic of the discussion. In these models, we control for the instructor (not shown). There is at least one significant treatment effect in six of the eight topics (exceptions are Civil Rights and Public Opinion). Consider the Campaigns & Elections topic, which has a significant treatment effect in three rubric categories (CLC, WQ, and RPT). Specifically, controlling for the course instructor, students in the treatment group receive 3.34 points more than students in the control group in the CLC category (the treatment effects for the WQ and RPT categories are 1.88 and 2.49 points, respectively). While we expect positive coefficients, there are several instances where the treatment has a negative effect, most notably for the Constitution board (six of the seven models show a negative coefficient). Here, students in the treatment group receive a final grade that is nearly 3.5 percentage points lower than those students in the control group.

Table 6. Priming frames and rubric scores for student success, by topic for all instructors.

	Discussion topic			
Rubric category (DV)	Campaigns and elections	Civil rights	Congress	Constitution
IAA				−3.84***
CCM				
OES				−2.64*
CLC	3.34**		2.25*	−3.29**
WQ	1.88*			−2.26*
RPT	2.49*			−5.75***
Grade				−3.39***

	Discussion topic			
Rubric category (DV)	Media	Political parties	Presidency	Public opinion
IAA				
CCM				
OES			−3.08**	
CLC	3.13*			
WQ	2.22*	−2.09*		
RPT				
Grade				

Note: Each entry summarizes a different model where the rubric category (DV) is regressed on the treatment condition as well as controls for the instructor (not shown). We run each of the seven models for each discussion board topic. Table only shows entries for significant coefficients (one tail tests) where: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Rubric Category Abbreviations: IAA: Ideas, Arguments, and Analysis; CCM: Connection to Course Material; OES: Outside Evidence and Support; CLC: Contribution to the Learning Community; WQ: Writing Quality; RPT: Required Postings/Timeliness.

The mixed results from the last two tables suggest there are unique features of each discussion topic as well as some differences in the grading by the three instructors which may mask treatment effects when we pool topics and/or instructors together. Therefore, we run the models for the seven outcome variables and eight topics separately by instructor. While a summary of these models with coefficients is available by instructor in the Supplemental Appendix (Tables A2, A3, and A4), we present a summary of these results for all instructors combined in Table 7, which shows the direction and level of significance for the treatment variable.[31] Each topic has at least two different significant treatment effects across the three instructors. The discussion board that most accurately and consistently matches our expectations is the Media topic. Between two of the instructors, six of the seven outcome variables show a significant treatment effect. In other words, for the Media topic in Instructor B's sections, students in the treatment group receive higher scores in all categories (save CCM and WQ) than students in the control group. In Instructor C's sections, there is a positive treatment effect for students in the WQ category. There are no significant treatment effects in the Media category for Instructor A. In some cases, however, the direction of the treatment effect varies by instructor. For example, Instructors B and C find opposite treatment effects for several of the outcome variables for the Constitution board.

Tables 6 and 7 both suggest the treatment frame can impact the rubric scores, depending on the instructor and topic, however, the results are not consistent. Sometimes, the treatment has no effect, sometimes it results in an improvement in scores (by as much as 8.24 percentage points), and other times it leads to a reduction in scores (by as much as 6.77 percentage points). There is little doubt that the topic and the framing of the prompt has the potential to affect the quality of the content in these

Table 7. Summary of regressions for rubric scores, by topic and instructor.

	Discussion topic											
	Campaigns and elections			Civil rights			Congress			Constitution		
	A	B	C	A	B	C	A	B	C	A	B	C
IAA			+*					+**				−***
CCM	−*		+*	+***	−*							−**
OES											+**	−***
CLC			+**						+*		+*	−***
WQ	+***	−**									+*	−**
RPT	+*										−**	−***
Grade			+*									−***

	Discussion Topic											
	Media			Political Parties			Presidency			Public Opinion		
	A	B	C	A	B	C	A	B	C	A	B	C
IAA		+**									−*	
CCM				+*								
OES		+**						−*	−*		−*	
CLC		+**										
WQ			+*			−**						
RPT		+**		+*								
Grade		+**		+*							−*	

Note: Each model is regressed on the treatment condition for each instructor (A, B, and C). Table only shows signs and level of significance where: *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.
Rubric Category Abbreviations: IAA: Ideas, Arguments, and Analysis; CCM: Connection to Course Material; OES: Outside Evidence and Support; CLC: Contribution to the Learning Community; WQ: Writing Quality; RPT: Required Postings/Timeliness.

discussion boards; as such, instructors must think carefully when designing these assignments to maximize their effectiveness.

Conclusion and discussion

Utilizing a 2×2 post-test experimental design across eight discussion topics in Introduction to American Government classes, we find mixed results for our hypotheses. We expected that the *discussion board topics* will impact student engagement (H1), content (H2), and quality (H3). Similarly, we expected *treatment conditions that include priming prompts* will lead to greater student engagement (H4), different content (H5), and higher quality responses (H6) compared to students in discussion groups with prompts that are more traditional.

Student engagement varies minimally across topics (H1) and between experimental conditions (H4). Students typically posted about 3 times per topic (which was the minimum number required to complete the assignment), and they write about 165 words over the course of 7 sentences. We did uncover a notable difference between the high levels of engagement in initial posts compared to the lower engagement when replying to classmates. Instructors should encourage students to sustain dialogue with their peers by critically analyzing the posts rather than simply agreeing with their classmates and meeting minimum requirements.

In some cases, the substantive topics (H2) and framing of prompts (H5) can impact the content of discussion boards, as shown visually in word clouds. Prompts that include priming generate more negative emotions from students in nearly all of our topics, but in only a few instances is this difference statistically significant. On the question of negative affect, scholars suggest this may be beneficial for students as negative affect is associated with better information recall (e.g., Marcus, Neuman, and MacKuen 2000). Furthermore, we know from the word clouds that students will use the words and phrases in the prompt, and that there can be important differences in the way students approach their responses to different topics (H2) as well as how they address traditional versus primed prompts (H5).

Finally, we assessed the quality of the discussion across topics (H3) and experimental frames (H6). We show that students use more sources and evidence as the course progresses, and—in terms of earning higher scores across our other learning outcomes—the treatment prompts that include priming have the potential to affect scores but not in a uniform manner. In sum, this research shows that instructors should be very mindful when designing discussion board assignments and thoughtful in the framing of the posts that start these conversations. Importantly, instructors should consider that the intent of the framing language may be inconsistent with how students interpret the prompt. For example, a prompt we might consider provocative could be interpreted by students as more leading, or simply more negative. Beyond these findings from our analyses, we think this project generates some suggestions for best practices in political science pedagogy research, using online discussion boards, and teaching Introduction to American Government.

For political science pedagogy research, using experiments of this kind is both low-cost and can help students meet the learning outcomes for classes. Prior to this study,

using online discussion boards was not normally part of our Introduction to American Government classes. However, by using eight relatively low-stakes assignments, we were able to give students more routinized feedback over the course of the term, shift some of the course grade burden from high-stakes exams, and have students engage with the material out of class in ways they might not have done otherwise. Moving forward, all instructors plan to continue to utilize discussion board assignments even after our study concludes due to their pedagogical value and the insight gleaned from this project. A similar design protocol can be done with writing assignments, group work, or any number of other assignments. In short, creating experimental conditions for pedagogy research can create a richer course experience for students.

The results of this project suggest, once again, that rubrics are vital for assessing assignments of this nature. For example—as evidenced in the rubric scores—the continual reinforcement to use evidence from *outside* of the course resulted in an increase with each iteration of the discussion board assignments. Still, there are ways to improve the use of rubrics. For instance, instructors and/or researchers can use a strategy session prior to the launch of the assignment to discuss rubric categories, how they are interpreted, etc., which would increase inter-coder reliability. Additionally, readability statistics may help validate rubric scores for "Writing Quality" and "Required Posts and Timeliness" (see our readability statistics presented in Tables A5 and A6 in the Supplemental Appendix).[32]

Finally, from the qualitative analysis and use of word clouds, we see that prompt content affects the vocabulary that students use in their responses. If the language of political science is important for students to learn—and we think it is—faculty need to model this in the assignment prompts they create. While this may be a simple mimicry device at the outset, as students use these words and phrases more (e.g., "polarization," "Duverger's Law," etc.) they will both learn what they mean and become more comfortable using them. For our part, we think it is crucial for instructors to use introductory courses to promote civil discourse and help students become more conversant in the language and approach of political science, which will improve learning outcomes in the short term and produce more well-rounded citizens in the long run.

Notes

1. Student behaviors and submissions with online discussion boards are often exercises in satisficing. By "meaningful, content-rich" discussions we mean evidence of engagement and consideration beyond a superficial "I agree with…" statement or some other banal minimalist behavior that is clearly instrumental in nature.
2. Online discussions spaced over the course of the semester allow for several relatively low-stakes assignments (compared to exams or term papers) for students to engage with the course material. Collectively, these assignments comprised 15% of the final grade for each course. We hoped these discussions would simulate discussions students may have outside of the classroom, but we acknowledge that the small contribution each discussion forum makes to the final grade could have an impact on the behavior of the students.
3. As discussed below, these experimental prompts include language that primes information relevant to each discussion topic. We expect students will interpret these frames differently from the students who receive the traditional prompts.

4. We chose our three specific outcomes of interest based on our collective goals for our students in introductory courses, which were informed by learning outcomes set by our departments and universities. We believe a similar research approach could inform instructors, regardless of their unique learning outcomes.
5. Researchers have even considered students' perceptions of learning through the integration of F2F and online discussions (Bliuc et al. 2010).
6. Of course, "best practices" will strongly correlate with student achievement because those students exhibiting behaviors we call best practices are likely the same students who earn the best scores.
7. We would like to thank several anonymous reviewers for helping us think more critically about how we conceptualize the framing of our discussion prompts. In earlier versions of this research, we labeled some prompts as "provocative," however, sometimes the language in our prompts could be read as leading, politically charged, or even activating a negative view toward the topic at hand. By classifying our experimental frames as "priming prompts," we acknowledge that the interpretation of our frames may vary by audience, the substantive topic of the discussion, as well as the specific language intended to prime the reader. What our treatment prompts do, fundamentally, is bring to front-of-mind something specific that is likely to generate affective engagement.
8. Theoretically, each student should participate in four control conditions and four treatment conditions over the course of the semester; furthermore, the network of students participating within each board should be different for each topic and experimental condition.
9. Indeed, this could be the case for any president. However, as Donald Trump was in office during our study period, we use his name as the more politically charged reference.
10. Table 1 presents the prompts and experimental conditions we used for our research; however, it is not our intent to suggest these are the ideal or only prompts to use for an introductory course in American politics. Indeed, our decisions on the language of our priming frames will impact the discussion to follow. As an anonymous reviewer rightly noted, using language such as "minorities" or "low income voters" may provoke implicit bias. Alternatively, the discussion may take a different tone if we had primed with terms like "voter equity," or "social and economic justice." Future research and practical application should always analyze the impacts of specific choices in priming language, perhaps through the use of additional experimental conditions.
11. While questions of external validity are fair in any experimental setting, we took several steps to improve the quality of our inferences. At the start of the semester, instructors informed students that discussion board posts would be utilized as part of an ongoing research project. Students reviewed and signed an IRB-approved consent form (Institution X Study ID: 17-0162; Institution Y: Human Subjects Review (HSR) approval 01/13/2017); however, they were not required to participate in the study (though they would still need to complete the course discussion board assignments). To assist with external validity, after this initial discussion the research study was not mentioned again, nor was the fact that students were assigned different prompts for each topic. As such, we hoped students would not think of the discussion as anything out of the ordinary for a college course.
12. At the conclusion of each discussion, we copied information from the LMS into a database, including the text of the post itself, the author, topic, timestamp, whether the post was an initial response to the prompt or a reply to one of their peers, and information about the specific course (e.g., semester, instructor, university, etc.).
13. Institution X utilizes Blackboard (http://www.blackboard.com/) while Institution Y uses Desire 2 Learn (https://www.d2l.com/).
14. To help account for the potential differences in the educational context surrounding each assignment, in the analysis to follow, we often disaggregate results by instructor or include appropriate controls in our regressions. In the end, we believe the value added by our

broader sample outweighs the potential for differences in instruction across faculty members.

15. The minimum number of posts required to complete each assignment was 3 (1 initial post and 2 replies to classmates). Many students posted more than this minimum threshold for completion; however, a number of students did not submit two replies to their classmates.
16. For reference, an average page of standard text (double spaced, 1" margins, 12pt Times New Roman) contains approximately 250 words. Instructors did not give students a specific length requirement for initial posts or replies.
17. Before building word clouds, we take a number of steps to clean the data. First, we use a package adept at text mining (*tm)* (Feinerer et al. 2018) to remove punctuation, capitalization, 'stopwords' (e.g., a, and, the, etc.), and extra white space. In addition, we use the *SnowballC* (Bouchet-Valat 2019) package to perform word stemming, which collapses words to their root for easier text analysis (e.g., "politics," "political," and "politician" each become "polit").
18. Here, we asked students to consider what the founding fathers might think of the state of the country today. In the treatment condition, we primed the "current partisan political climate and the Trump presidency."
19. Students mentioned "Trump" just 13 times in the control group but 317 times in the treatment condition. The stem "parti-" (e.g., partisan, parties) was used 76 times in the control group and 172 times in the treatment group. Students brought up immigration nearly 4 times as often in the treatment group (98) compared to the control group (25).
20. One of the most common words used in both groups is "constitut-," which was used 458 times in the control group and 351 times in the treatment group.
21. See the Figure A1 in the Appendix for comparison word clouds for the other topics.
22. While there are several R packages for these methods, here we utilized *tidytext* (Queiroz et al. 2019).
23. For additional information on this lexicon, please visit: http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm.
24. This lexicon categorizes some words within multiple emotions/sentiments. For example, "abandon" evokes both fear and sadness but also reveals a negative sentiment.
25. For example, in the Campaigns & Elections board, there are 6,678 negative words and 6,905 positive words. As percentages of the total negative or positive words, this results in a Net Sentiment value of 1.6% (50.8% positive—49.2% negative), or a sentiment that is just slightly more positive than negative.
26. As an anonymous reviewer pointed out, this finding may be a function of students in an introductory course. In future research, it would be interesting to compare the Net Sentiment using a similar experimental design in an upper level seminar where students have greater grasp of the discipline.
27. VALUE rubrics are available from the AAC&U website (https://www.aacu.org/value/rubrics).
28. Results by instructor and by topic available upon request.
29. These variables also control for temporal grading variation for each instructor. We acknowledge, for example, that instructors may consciously or subconsciously grade differently at the start of the semester.
30. Full results available upon request.
31. Full results available upon request.
32. We conducted a preliminary readability analysis by topic, treatment condition, initial post vs. reply posts, and by instructor. Using the R package *quanteda* (Quantitative Analysis of Textual Data) (Benoit et al. 2018), we present the average number of words and sentences in each post in addition to five different readability measures, each of which returns a value that corresponds to the approximate grade level of the text.

Acknowledgments

Previous iterations of this paper were presented at the 2017 and 2019 American Political Science Association annual meetings in San Francisco and Washington, DC, respectively. We also thank Justin Brantley for his assistance compiling data.

Notes on contributors

Aaron S. King is an associate professor and the coordinator of the undergraduate program in political science within the Department of Public & International Affairs at UNC Wilmington. His research and teaching focuses on several topics within American politics, including political institutions, ambition, representation, and competition.

J. Benjamin Taylor is an assistant professor of political science in the School of Government & International Affairs at Kennesaw State University. His research and teaching focuses on American politics with a specialty in political behavior and political communication.

Brian M. Webb is an associate professor of political science at Gordon State College. His research and teaching focuses on American politics broadly, but with a particular focus on political institutions, Congressional procedure, and Georgia politics.

ORCID

Aaron S. King http://orcid.org/0000-0002-2435-0449
J. Benjamin Taylor http://orcid.org/0000-0002-8541-3466

References

Backer, David. 2016. "The Purpose of Online Discussion." *Hybrid Pedagogy: A Digital Journal of Learning, Teaching, and Technology*. http://www.digitalpedagogylab.com/hybridped/purpose-online-discussion/ (August 4, 2017).

Barber, Benjamin R. 1999. "The discourse of civility." In *Citizen Competence and Democratic Institutions*, eds. Stephen L. Elkin and Karol Edward Soltan. University Park, PA: The Pennsylvania State University Press, 39–48.

Barber, Benjamin R. 2003. *Strong Democracy: Participatory Politics for a New Age*. Berkeley, CA: University of California Press.

Barberá, Pablo. 2015. "Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data." *Political Analysis* 23 (1):76–91. doi: 10.1093/pan/mpu011.

Bartels, Larry M. 2005. "Homer Gets a Tax Cut: Inequality and Public Policy in the American Mind." *Perspectives on Politics* 3 (01):15–31. doi: 10.1017/S1537592705050036.

Bartels, Larry M. 2007. "Homer Gets a Warm Hug: A Note on Ignorance and Extenuation." *Perspectives on Politics* 5:785–790.

Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo. 2018. "Quanteda: An R Package for the Quantitative Analysis of Textual Data." *Journal of Open Source Software* 3 (30):774. doi: 10.21105/joss.00774.

Bliuc, Ana-Maria, Robert Ellis, Peter Goodyear, and Leanne Piggott. 2010. "Learning through Face-to-Face and Online Discussions: Associations between Students' Conceptions, Approaches and Academic Performance in Political Science." *British Journal of Educational Technology* 41 (3):512–524. doi: 10.1111/j.1467-8535.2009.00966.x.

Bolsen, Toby. 2013. "A Light Bulb Goes on: Norms, Rhetoric, and Actions for the Public Good." *Political Behavior* 35 (1):1–20. doi: 10.1007/s11109-011-9186-5.

Bonwell, C. C., and James A. Eison. 1991. *Active Learning: Creating Excitement in the Classroom*. Washington, DC: ERIC Clearinghouse on Higher Education. https://eric.ed.gov/?id=ED336049 (July 1, 2020).

Bouchet-Valat, Milan. 2019. "SnowballC: Snowball Stemmers Based on the C "libstemmer" UTF-8 Library." https://CRAN.R-project.org/package=SnowballC (July 31, 2019).

Brookfield, Stephen D., and Stephen Preskill. 2012. *Discussion as a Way of Teaching: Tools and Techniques for Democratic Classrooms*. New York, NY: John Wiley & Sons.

Coe, Kevin, Kate Kenski, and Stephen A. Rains. 2014. "Online and Uncivil? Patterns and Determinants of Incivility in Newspaper Website Comments." *Journal of Communication* 64 (4):658–679. doi: 10.1111/jcom.12104.

Dabbagh, Nada, and Brenda Bannan-Ritland. 2005. *Online Learning: Concepts, Strategies, and Application*. New York, NY: Prentice Hall.

Davies, Jo, and Martin Graff. 2005. "Performance in E-Learning: Online Participation and Student Grades." *British Journal of Educational Technology* 36 (4):657–663. doi: 10.1111/j.1467-8535.2005.00542.x.

Davis, Meredith. 2017. *Teaching Design: a Guide to Curriculum and Pedagogy for College Design Faculty and Teachers Who Use Design in Their Classrooms*. New York, NY: Allworth Press.

Dilliplane, Susanna. 2011. "All the News You Want to Hear: The Impact of Partisan News Exposure on Political Participation." *Public Opinion Quarterly* 75 (2):287–316. doi: 10.1093/poq/nfr006.

Druckman, James N. 2011. "What's It All About?: Framing in Political Science." In *Perspectives on Framing*, ed. Gideon Keren. New York, NY: Psychology Press, 279–302.

Entman, Robert M. 2004. *Projections of Power: Framing News, Public Opinion, and U.S. Foreign Policy*. Chicago, IL: University of Chicago Press.

Feinerer, Ingo, and Kurt Hornik. 2018. "Tm: Text Mining Package." https://CRAN.R-project.org/package=tm (July 31, 2019).

Fellows, Ian. 2018. "Wordcloud: Word Clouds." https://CRAN.R-project.org/package=wordcloud (July 31, 2019).

Fishkin, James S. 1991. *Democracy and Deliberation: New Directions for Democratic Reform*. New Haven, CT: Yale University Press.

Glazier, Rebecca A. 2016. "Building Rapport to Improve Retention and Success in Online Classes." *Journal of Political Science Education* 12 (4):437–456. doi: 10.1080/15512169.2016.1155994.

Hamann, Kerstin, Philip H. Pollock, and Bruce M. Wilson. 2009. "Learning from 'Listening' to Peers in Online Political Science Classes." *Journal of Political Science Education* 5 (1):1–11. doi: 10.1080/15512160802612011.

Hamann, Kerstin, Philip H. Pollock, and Bruce M. Wilson. 2012. "Assessing Student Perceptions of the Benefits of Discussions in Small-Group, Large-Class, and Online Learning Contexts." *College Teaching* 60 (2):65–75. doi: 10.1080/87567555.2011.633407.

Hara, Noriko, Curtis Jay Bonk, and Charoula Angeli. 2000. "Content Analysis of Online Discussion in an Applied Educational Psychology Course." *Instructional Science* 28 (2): 115–152. doi: 10.1023/A:1003764722829.

Homburg, Christian, Laura Ehm, and Martin Artz. 2015. "Measuring and Managing Consumer Sentiment in an Online Community Environment." *Journal of Marketing Research* 52 (5): 629–641. doi: 10.1509/jmr.11.0448.

Immordino-Yang, Mary Helen, and Antonio Damasio. 2007. "We Feel, Therefore We Learn: The Relevance of Affective and Social Neuroscience to Education." *Mind, Brain, and Education* 1 (1):3–10. doi: 10.1111/j.1751-228X.2007.00004.x.

Janssens-Bevernage, Anouk. 2014. "'Post Once, Reply Twice': Uninspiring Online Discussion Boards and What to Do about Them." *DynaMind eLearning*. https://dynamind-elearning.com/2014/03/post-once-reply-twice-uninspiring-online-discussion-boards-and-what-to-do-about-them/ (August 4, 2017).

Jones, E. L., and Ronald C. Jones. 2014. "The Online Discussion Board: Opening the Gateway to New Learning." *Faculty Focus | Higher Ed Teaching & Learning*. https://www.facultyfocus.com/

articles/online-education/online-discussion-forum-opening-gateway-new-learning/ (July 29, 2017).
Kahneman, Daniel, and Amos Tversky. 2000. *Choices, Values, and Frames*. New York, NY: Cambridge University Press.
King, Aaron S., Frank J. Orlando, and David B. Sparks. 2016. "Ideological Extremity and Success in Primary Elections: Drawing Inferences from the Twitter Network." *Social Science Computer Review* 34 (4):395–415. doi: 10.1177/0894439315595483.
Klar, Samara. 2014. "Partisanship in a Social Setting." *American Journal of Political Science* 58 (3):687–704. doi: 10.1111/ajps.12087.
Krentler, Kathleen A., and Laura A. Willis-Flurry. 2005. "Does Technology Enhance Actual Student Learning? The Case of Online Discussion Boards." *Journal of Education for Business* 80 (6):316–321. doi: 10.3200/JOEB.80.6.316-321.
Levendusky, Matthew. 2013. *How Partisan Media Polarize America*. Chicago, IL: University of Chicago Press.
Marcus, George E., W. R. Neuman, and MacKuen Michael. 2000. *Affective Intelligence and Political Judgment*. Chicago, IL: University of Chicago Press.
Marcus, George E. 2002. *The Sentimental Citizen: Emotion in Democratic Politics*. University Park, PA: Penn State Press.
Marcus, George E., John L. Sullivan, Elizabeth Theiss-Morse, and Daniel Stevens. 2005. "The Emotional Foundation of Political Cognition: The Impact of Extrinsic Anxiety on the Formation of Political Tolerance Judgments." *Political Psychology* 26 (6):949–963. doi: 10.1111/j.1467-9221.2005.00452.x.
McGuire, Beverley Foulks. 2016. "Integrating the Intangibles into Asynchronous Online Instruction: Strategies for Improving Interaction and Social Presence." *Journal of Teaching Effectiveness* 16 (3):62–75.
Mohammad, Saif M., and Peter D. Turney. 2013. "Crowdsourcing a Word–Emotion Association Lexicon." *Computational Intelligence* 29 (3):436–465. doi: 10.1111/j.1467-8640.2012.00460.x.
Morris, John Michael, and Jesse Stommel. 2013. "The Discussion Forum Is Dead; Long Live the Discussion Forum." *Hybrid Pedagogy*. http://www.digitalpedagogylab.com/hybridped/the-discussion-forum-is-dead-long-live-the-discussion-forum/ (August 4, 2017).
Muddiman, Ashley, and Natalie Jomini Stroud. 2017. "News Values, Cognitive Biases, and Partisan Incivility in Comment Sections." *Journal of Communication* 67 (4):586–609. doi: 10.1111/jcom.12312.
Mutz, Diana C. 2006. *Hearing the Other Side: Deliberative versus Participatory Democracy*. Cambridge, UK: Cambridge University Press.
Queiroz, Gabriela De., et al. 2019. "Tidytext: Text Mining Using "Dplyr", "Ggplot2", and Other Tidy Tools." https://CRAN.R-project.org/package=tidytext (July 31, 2019).
Reddy, Y. Malini, and Heidi Andrade. 2010. "A Review of Rubric Use in Higher Education." *Assessment & Evaluation in Higher Education* 35 (4):435–448. doi: 10.1080/02602930902862859.
Rovai, Alfred P. 2007. "Facilitating Online Discussions Effectively." *The Internet and Higher Education* 10 (1):77–88. doi: 10.1016/j.iheduc.2006.10.001.
Stanley, J. Woody, and Christopher Weare. 2004. "The Effects of Internet Use on Political Participation: Evidence from an Agency Online Discussion Forum." *Administration & Society* 36 (5):503–527. doi: 10.1177/0095399704268503.
Stevens, Danielle D., and Antonia J. Levi. 2013. *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning*. 2nd ed. Sterling, VA: Stylus Publishing.
Taylor, J. Benjamin. 2017. *Extreme Media and American Politics: In Defense of Extremity*. New York, NY: Palgrave Macmillan. http://www.palgrave.com/us/book/9783319411828 (March 22, 2017).
Williams, Leonard, and Mary Lahman. 2011. "Online Discussion, Student Engagement, and Critical Thinking." *Journal of Political Science Education* 7 (2):143–162. doi: 10.1080/15512169.2011.564919.

Wilson, Bruce M., Philip H. Pollock, and Kerstin Hamann. 2007. "Does Active Learning Enhance Learner Outcomes? Evidence from Discussion Participation in Online Classes." *Journal of Political Science Education* 3 (2):131–142. doi: 10.1080/15512160701338304.

Wilson, Bruce M., Philip H. Pollock, and Kerstin Hamann. 2014. "Spillover Effects in Online Discussions: Assessing the Effectiveness of Student Preceptors." *Journal of Information Technology & Politics* 11 (1):15–24.

Wu, Dezhi, and Starr Roxanne Hiltz. 2004. "Predicting Learning from Asynchronous Online Discussion." *Journal of Asynchronous Learning Networks* 8 (2):139–152.

Young, Lori, and Stuart Soroka. 2012. "Affective News: The Automated Coding of Sentiment in Political Texts." *Political Communication* 29 (2):205–231. doi: 10.1080/10584609.2012.671234.