

11-6-2017

Prompting Deliberation about Nanotechnology: Information, Instruction, and Discussion Effects on Individual Engagement and Knowledge

Lisa M. PytlikZillig
University of Nebraska, lpytlikz@nebraska.edu

Myiah J. Hutchens
Washington State University, myiah.hutchens@gmail.com

Peter Muhlberger
University of Nebraska, pmuhl1848@gmail.com

Alan J. Tomkins
University of Nebraska, atomkins@nebraska.edu

Follow this and additional works at: <https://www.publicdeliberation.net/jpd>

 Part of the [Civic and Community Engagement Commons](#), [Cognitive Psychology Commons](#), [Education Commons](#), [Gender, Race, Sexuality, and Ethnicity in Communication Commons](#), [Personality and Social Contexts Commons](#), [Political Science Commons](#), and the [Social Psychology Commons](#)

Recommended Citation

PytlikZillig, Lisa M.; Hutchens, Myiah J.; Muhlberger, Peter; and Tomkins, Alan J. (2017) "Prompting Deliberation about Nanotechnology: Information, Instruction, and Discussion Effects on Individual Engagement and Knowledge," *Journal of Public Deliberation*: Vol. 13 : Iss. 2 , Article 2.

Available at: <https://www.publicdeliberation.net/jpd/vol13/iss2/art2>

This Article is brought to you for free and open access by Public Deliberation. It has been accepted for inclusion in Journal of Public Deliberation by an authorized editor of Public Deliberation.

Prompting Deliberation about Nanotechnology: Information, Instruction, and Discussion Effects on Individual Engagement and Knowledge

Abstract

Deliberative (and educational) theories typically predict knowledge gains will be enhanced by information structure and discussion. In two studies, we experimentally manipulated key features of deliberative public engagement (information, instructions, and discussion) and measured impacts on cognitive-affective engagement and knowledge about nanotechnology. We also examined the direct and moderating impacts of individual differences in need for cognition and gender. Findings indicated little impact of information (organized by topic or by pro-con relevance). Instructions (prompts to think critically) decreased engagement in Study 1, and increased it in Study 2, but did not impact post-knowledge. Group discussion had strong positive benefits for self-reported cognitive-affective engagement across studies. Also, for some types of engagement, effects were more positive for women than men. When predicting knowledge, there also was some evidence that discussion was more positive for women than men. Finally, need for cognition positively predicted engagement and knowledge gains, but rarely moderated the experimental effects. Given these mixed results, future research should continue to test theoretical assumptions about the effects of specific deliberative design features.

Keywords

Public Deliberation, deliberation prompts, knowledge, critical thinking, cognitive engagement, learning, affective engagement, gender differences, discussion, information organization, need for cognition

Acknowledgements

The authors would like to thank University of Nebraska-Lincoln Professors Joseph A. Turner and Yuris Dzenis for help creating the background document, and UNL graduate students Jayme Neiman, Frank Gonzalez, Ryan Anderson, and numerous REU students for assistance with the studies. This research was funded by the National Science Foundation (0965465), with supplemental funds provided for Research Education for Undergraduates (REU). The opinions expressed are the authors and not the opinions of the NSF.

Introduction

There are many desirable potential outcomes of participating in public engagements. Learning outcomes are especially important because knowledge is a prerequisite to offering *informed* policy input, which may make the input more useful and influential (Guston, 2014; Muhlberger & Weber, 2006). Prior research suggests deliberative public engagements, in particular, may improve public understanding of science and technology by providing participants with opportunities to study relevant information as they form their preferences (e.g., Farrar et al., 2010). However, not all studies find positive effects of deliberation (Delli Carpini, Cook, & Jacobs, 2004; Ryfe, 2005), and even when effects are found, it is difficult for researchers to identify the mechanisms responsible (e.g., Sanders, 2012).

Experiments investigating the effects of specific features of public engagement are especially important for advancing theoretical understanding of what features of public engagements work for what purposes and why, and to guide the design of effective engagements (PytlikZillig & Tomkins, 2011). In addition, because of concerns relating to issues of equality and engagement (Benhabib, 2002), it is important to examine potential moderators. Not all publics have equal information or influence relating to political or policy issues, and little research has examined whether certain deliberative mechanisms favor some groups over others (Fraile, 2014; Hickerson & Gastil, 2008; Karpowitz, Mendelberg, & Shaker, 2012).

Deliberative engagements include features such as provision of balanced information, encouragement of deep cognitive engagement, and group discussion (Fishkin & Luskin, 2005). Theory suggests these features may promote increased knowledge and potentially more well-justified attitudes and policy preferences (Chambers, 2003; Mendelberg, 2002). However, there are numerous empirical gaps in these theorized connections. For example, despite the centrality of deep cognitive engagement to deliberative theory, few studies of deliberative practice explicitly measure cognitive engagement, or the variety of other ways people may engage. Even fewer attempt to causally connect different forms of individual engagement to specific deliberative design features and outcomes, such as increased knowledge or understanding.

To begin to fill this gap, in the present studies, we experimentally varied features of deliberation (information, instructions, and discussion), and measured the individual and combined impacts of these features on individual-level engagement and knowledge. Further, we examined potential moderation by two other variables: gender—which is a longstanding basis of political inequality

(Benhabib, 2002)—and individual differences in need for cognition (the tendency to enjoy and use effortful and deep thinking processes (Cacioppo, Petty, Feinstein, & Jarvis, 1996)—a variable especially relevant to deliberation.

We conducted our studies in the context of engaging college science students in deliberations about potential ethical, legal, and social implications (ELSI) associated with nanotechnology. While the college classroom context is not representative of the majority of public engagement contexts, it is one such context, and one that facilitates controlled experimentation. In addition, findings from studies of the design of deliberative discussions in this context can specifically improve the use of deliberative practices when helping students consider ELSI implications of new science and technology developments—a practice which is increasingly encouraged (Barsoum, Sellers, Campbell, Heyer, & Paradise, 2013). Finally, findings in this context may suggest possibilities that should be investigated in other public engagement contexts.

Background

The Theoretical Importance of Varieties of Individual Engagement

There are a large number of ways that individuals might “engage” during public engagement activities (PytlikZillig, Hutchens, Muhlberger, Wang, Harris, Neiman, & Tomkins, 2013), but one type—deep cognitive engagement—defines deliberation (Mercier & Landemore, 2012; Morrell, 2005). Psychologists distinguish between deep, effortful, controlled (type 2) versus automatic, surface (type 1) cognitive processes (Chaiken, 1980; Kahneman, 2011). In educational psychology, surface processing refers to simple acts such as reading and repetition, whereas deep processing refers to active and metacognitive activities that promote the integration of old and new knowledge—for example, questioning, elaborating, and restructuring one’s understandings (Chin & Brown, 2000). Prior research suggests deep cognitive processing can create larger and longer-lasting knowledge gains (Dinsmore & Alexander, 2012; Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013), providing support for encouraging deep processing during public engagements.

Beyond deep cognitive engagement, a complete picture of deliberation should also consider how people emotionally, behaviorally, and socially engage (Fredricks, Blumenfeld, & Paris, 2004). A number of deliberative theorists refer to the importance of open-minded, conscientious, and empathetic engagement during deliberative activities (Fishkin & Luskin, 2005; Mendelberg, 2002). Furthermore, critics of the “deliberative ideal” argue against emphasizing or

privileging certain rational processes to the exclusion of emotions, noting that such emphasis may undermine input from certain groups (Benhabib, 2002; Hickerson & Gastil, 2008; Martin, 2012).

Despite the importance of “how” people engage during public engagement, few studies of public engagement have assessed individual-level cognitive, emotional, and/or behavioral engagement. These studies nonetheless point to the importance of different ways of engaging. Warnick, Xenos, Endres, and Gastil (2005), for example, experimentally manipulated two forms of interactivity with web-based political information and found that while both forms increased cognitive engagement when used alone, simultaneous use had a negative effect, suggesting there is a such thing as too much engagement encouragement. Also, in political contexts, Marcus, MacKuen, and colleagues demonstrate that different emotions are associated with different information-seeking behaviors (MacKuen, Wolak, Keele, & Marcus, 2010; Marcus, Neuman, & MacKuen, 2000). In the context of deliberative engagements, however, while many studies have examined knowledge gains, and a few have examined connections between specific features of deliberation and knowledge or attitude changes (e.g., Farrar et al., 2010; Morrell, 2005; Muhlberger & Weber, 2006), prior studies have not directly investigated the role of different forms of cognitive-affective engagement in increasing knowledge. Our studies begin to fill this gap.

Deliberative Design Features

Cognitive, affective, and behavioral engagement is frequently assessed in educational environments (Fredricks et al., 2004), providing bases for hypotheses about the effects of certain deliberative features. In the present studies, we examined the impacts of three design features commonly used in deliberative engagements: information features, instructions prompting deep cognitive processing, and discussion.

Information Organization. Public deliberation practitioners commonly argue that balanced information is an important component of effective deliberation (Burkhalter, Gastil, & Kelshaw, 2002; Lukensmeyer & Torres, 2006). However, there are various ways in which that information might be structured. We hypothesized that text organized into *pros and cons* would facilitate comparing different perspectives and enhance learning relative to text organized by *topics*. Our hypothesis stemmed from research showing that students who take notes in matrix formats (with columns and rows encouraging comparison and contrast), tend to learn more than those who take outline or linear notes (Robinson & Kiewra, 1995). Furthermore, texts structured in a compare-contrast form tend to

be associated with improved recall, compared to texts organized linearly or descriptively (Bohn-Gettler & Kendeou, 2014).

Promoting Deep Engagement. In deliberative engagements, the goal of deep engagement is promoted by instructing participants to share diverse ideas and engage in “reasoned argument” (McCoy & Scully, 2002, p. 124). We hypothesized that prompting participants to think critically (versus simply to attend to the information) would increase deep cognitive engagement and learning. This hypothesis was based on research finding students employ different strategies depending on their goals. For example, when reading for entertainment, students create associations that are not essential to understanding the text; but when studying, they make more explanatory connections and remember more of the information (Bohn-Gettler & Kendeou, 2014). Prompts used to encourage self-monitoring of one’s learning (Kauffman, Zhao, & Yang, 2011), or to explain, justify, and provide arguments for different points of view, also have been found to effectively enhance deep engagement, conceptual change, and reasoning (Chi, Leeuw, Chiu, & LaVancher, 1994; Garcia-Mila, Gilabert, Erduran, & Felton, 2013; Nussbaum & Kardash, 2005).

Discussion. A third, nearly ubiquitous, feature of deliberative engagements is group discussion. According to Burkhalter et al. (2002, p. 401), in addition to being characterized by “careful weighing” of different viewpoints, “deliberation is characterized by the performance of a set of communicative behaviors that promote thorough group discussion.” While people can deliberate alone (Goodin & Niemeyer, 2003), theorized benefits of group discussion date back to Lewin (e.g., 1943) and include exposure to and greater understanding of other viewpoints; and changing, clarifying, elaborating, legitimizing, and committing to viewpoints as participants explain them to each other and contribute to group decision-making. Here we focus on discussion’s potential beneficial impacts on participant learning. Because discussion has been found to have positive impacts on learning in other contexts (Levin, 1995; Van Blankenstein, Dolmans, Van der Vleuten, & Schmidt, 2013), we hypothesized it would enhance knowledge. In actuality, however, research examining the effects of discussion on knowledge in public engagement contexts has found only small effects (Mendelberg, 2002). Muhlberger and Weber (2006), for example, carefully separated the effects of reading from discussion and found learning occurred during reading, but no significant additional learning gains occurred during discussion. Their findings align with similar findings from deliberative studies examining attitude changes (Goodin & Niemeyer, 2003; Luskin, Fishkin, & Jowell, 2002). These counter-intuitive findings deserve additional attention. Thus, in addition to examining discussion effects on learning, we investigated potential moderators.

Potential Moderators

We examined three sources of moderation: moderation by other deliberative features, gender, and need for cognition. We examined moderation between different deliberative features because of prior research suggesting that there is an optimal amount of engagement promotion (Warnick et al., 2005). A negative interaction between features of information, instructions, and discussion would indicate an upper limit to how many positive features continue to increase engagement and learning. Alternatively, it is possible that features may build upon and amplify others. For example, positive impacts of a pro-con information organization might be accentuated through discussion with one's peers.

Second, although most theorists see deliberative, deep, strictly rational thinking methods as a beneficial ideal, the “difference critique” (Benhabib, 2002; Hickerson & Gastil, 2008) suggests deliberative methods may favor some groups more than others. Prior research finds that men and women communicate differently (Dow & Wood, 2006) and prefer different processes in different contexts (Karpowitz et al., 2012). Thus, it may be that encouraging deep, logical, critical-thinking processes is more beneficial for men than women. On the other hand, research finding gender equivalence in satisfaction with deliberative processes used by juries (Hickerson & Gastil, 2008) suggests deliberation is not inherently more negative for women. Also, Fraile (2014) finds evidence that deliberative activities may actually reduce gender gaps by providing women with information and improving their confidence expressing that information. Thus, it is important to examine whether particular deliberative features accentuate or reduce such differences.

Another relevant individual difference is need for cognition (NFC)—the tendency to enjoy and use effortful cognitive processing strategies (Cacioppo et al., 1996). Persons high in NFC have been identified as especially likely to participate in deliberations, be more resistant to the arguments of others, and have more influence (Delli Carpini et al., 2004). It may be beneficial to attempt to design deliberations in a way that helps to “level the playing field” of influence of persons high and low in NFC, or at least do not interact with NFC to widen the gap. However, it is not clear how different deliberative features might be moderated by NFC. For example, will instructions designed to increase deep processing (and knowledge) be more effective for participants high in NFC because they appreciate such instructions? Or will such instructions be most effective for participants low in NFC because they need reminders to think deeply, whereas high NFC persons deeply engage even without such prompts?

Methods¹

Participants

Participants in two studies (Study 1, S1: $n = 192$, 53% female; Study 2, S2: $n = 278^2$, 58% female) were enrolled in an introductory biology course for science majors at a large public university, and the majority of participants (73-75%) declared a science-related major. Self-reported political parties reflected a conservative leaning overall (41-45%, Republican, 24-26% Democrat, 29-35% Independent/other) and prior knowledge about nanotechnology was low, with the majority (72-86%) indicating being not at all or only slightly familiar with nanotechnology-related topics.

Procedures and Experimental Conditions

Both studies used a longitudinal design in which participants first completed a pre-survey assessing knowledge, demographics, and existing attitudes (Time 1, T1). Later (after 10 weeks in S1, after 1 week in S2), participants received a lecture on ethical, legal, and social issues related to science and saw a short video about nanotechnology. Immediately after the video, they were assigned online interactive homework readings composed of background information accompanied by experimentally manipulated instructions (Time 2, T2). The following week, they engaged in an in-class deliberation, experimentally manipulated to include or exclude discussion (Time 3, T3). During the deliberation, students were given access to online and paper versions of the background information they had been assigned to read as homework. Finally, during the week after the discussion, they completed a post-deliberation survey (Time 4, T4) as homework. Participation in the deliberative activities was a course requirement and students received course points for participation. However, their participation was not graded for quality and they were not required to allow the researchers to use their data for this study.

Reading, information organization, and instructions. The online, interactive background reading that students completed as homework drew from peer-reviewed sources, described ways in which nanotechnology is currently being used and its potential future applications, and included links to additional

¹ Copies of all materials and measures are available from the corresponding author.

² There were more than 300 students enrolled during S2, but two sections of students engaged online rather than face-to-face and thus were omitted from the present analyses.

information. In S1, the background information contained approximately 2,500 words, and participants were randomly assigned to one of three instructions while reading the background information. The *general engagement* condition asked them to “list five insights, realizations, reactions, or new things that you learned as a result of reading the background document or exploring the additional resources in that document.” The *information organization* condition asked them to list, from the reading, what some people claim are benefits and risks of nanogenomics research, and reasons for and against heavy regulation of such research, and then to rate the extent to which they agreed with the claims they listed. The *critical thinking* condition involved first describing to students an approach to critical analysis (Fulkerson, 1996), then asking participants to practice applying the approach to a sample problem, and finally to apply the approach to claims about nanotechnology’s risks and benefits and reasons for heavy or light regulation of nanogenomic research that they found in the reading.

Based on results from S1 (discussed below), a number of changes were made to S2. Because the S1 instruction conditions were not as effective as expected, in S2 we changed instructions for critical thinking and made them independent from information organization. S2 thus involved four reading conditions varied in a 2x2 design. The first factor varied instructions adapted from S1: *General engagement* prompts asked students to report what they found interesting or surprising about the reading; *critical thinking* prompts briefly defined aspects of critical thinking (e.g., characteristics of bias and quality sources) and simply asked students to apply critical-thinking skills to their reading (but without requiring initial practice as in S1). The second factor varied the information organization provided to the students. Additional background information was provided (length of S2 information was about 4,500 words). Information was either presented organized by *topics* (e.g., discussing risks and benefits, issues of autonomy, changes in society) or organized according to contrasting of *pro-con* perspectives, modeled after the National Issues Forum (providing descriptions of two, pro and con, perspectives and their action implications, as well as support for, trade-offs of, and opposition to each perspective). Although the structure was different, the same topics, issues, facts, and links to additional information were included in both versions of information.

Deliberation activities and social context. The in-class deliberation lasted 40-50 minutes, and all participants were given the same written descriptions of potential future scenarios to respond to (e.g., describing potential future use of nanogenomic research for cystic fibrosis prevention and improvement of human memory). Condition-appropriate background readings were also available in hard copy format or via online links provided to the students. During deliberation,

participants were randomly assigned to one of two social conditions. In the *group condition*, participants read the scenarios with two to four others from their same reading condition, discussing their reactions and opinions. Trained moderators (researchers and students who had been in the course and engaged in similar exercises the prior semester) led the discussion and instructed students to listen to and respond to one another but noted they were not required to come to consensus and should form their own opinions. All participants individually typed their reactions into a web-based form. Those assigned to the *individual condition* were in a separate quiet room, individually reacting to the same scenarios using the same web-based form, but without peer discussion. As previously noted, all students had access to condition-appropriate versions of the background information that they had read as homework.

Measured Variables

Varieties of engagement. Immediately after the reading (T2) and again after the deliberation activities (T3), students were asked to report their engagement by responding to items from the Varieties of Engagement (VIE) scales (PytlikZillig et al., 2013).³ The stem for all items was “during the assignment I...,” and responses used a 1 (not at all) to 5 (a great deal) scale. Measures of *conscientious* engagement (five items, e.g., felt focused, carefully evaluated the relevance of various arguments; Cronbach α s ranged from .80-.82) and *metacognitive/active learning* engagement (four to five items, e.g., tried to think about how the topics I was reading about related to other things I know, identified questions that I still had about the topics; α s = .77-.80) were expected to most reflect deep and effortful cognitive engagement. In addition, we assessed *social* (two to four items, e.g., discussed my ideas about the topics with others, asked others what they thought about the topics and issues; α = .88-.95), *closed-minded* (two to four items, e.g., felt... closed, like my mind was made up; S1 α s = .50-.52; S2 α s = .79-.81), *bored* (two to seven items, e.g., felt...like I wished I were doing something else, bored; α s = .80-.90), and *angry* (two to six items, felt...annoyed, frustrated; α s = .73-.92) engagement states. In S2, we also assessed *creative* (Five to six items, e.g., felt creative, used my imagination; α s =.85-.88) and *open-minded* (three items, e.g., felt open-minded, tried hard to understand perspectives that were different from mine; α s = .71-.74) engagement.

Need for cognition. Need for cognition (NFC) was assessed at T1 using seven items (e.g., “the notion of thinking abstractly is appealing to me,” α s = .80-.83) taken from the short version of the need for cognition scale (Cacioppo et al.,

³ Because the VIE scales were under development, the items used for each scale varied between studies 1 and 2.

1996). Participants rated their agreement (1=strongly disagree, 6=strongly agree) with the items, and items were reverse coded as needed to reflect high levels of NFC before averaging.

Knowledge. Knowledge measures comprised multiple-choice and true-false items (e.g., “how many nanometers are in 1 meter?”). Prior to analyses, individual items were screened for quality, and unclear or confusing items were omitted.⁴ In S1, knowledge was assessed with a random subset of three (out of six) knowledge questions assigned immediately prior to reading activities (pre-T2) and with all six pre-items plus five new items post all activities (T4). In S2, knowledge was assessed with five questions administered at both T1 and T4 and an additional 14 questions at T4, four of which were *new*, and 10 of which were *mixed* seen/unseen questions (students were assigned a random selection of five of these at other points in the study). Thus, the post-items always included some items that were *repeated* at T1 and T4, and some items that were *new* at T4 (and an additional category of *mixed* questions for S2). Because repeated knowledge questions are impacted by prior exposure (which primes attention for pre-test information), we created separate scores for repeated, new, and mixed items, and tested for differences using MANCOVA (using T1 knowledge as a covariate). At each time point, knowledge was computed as percent correct, except S1, pre-T2 items were transformed into z-scores and then averaged because students had received different item subsets.

Results

Preliminary Analyses and Analytic Strategy

At the end of the reading homework assignment, participants were asked to confidentially indicate if they had given honest answers to the questions (rather than, for example, answering randomly). Student responses to these questions were examined, and data points were excluded from relevant analyses if students indicated they mostly had not answered honestly (13 students were dropped from some of the analyses; for S1, $n=9$, and S2, $n=4$). We also examined the success of our random assignment procedures using univariate ANOVAs to compare experimental groups on a number of variables assessed at T1. Across studies, between-condition comparisons on gender, political party, ideology, interest in

⁴ Screening consisted of examining response patterns as well as examining item-total correlations and reconsidering item face validity in light of such statistics.

local or national politics, NFC, and political efficacy found only one significant association.⁵

To test our hypotheses, we used multivariate analyses of variance and covariance (MANCOVA), examining the effects of our experimental conditions, gender, and NFC on individual engagement and knowledge. After the reading condition at T2, we examined the impact of the reading conditions only. After the deliberation at T3, we examined the impacts of both the reading and discussion conditions. For each of these analyses, we examined assumptions for multivariate analyses, including normality of each variable, univariate and multivariate outliers, and Box's M to test for homoscedasticity (using $p < .005$ for significance level, as suggested by Huberty & Petoskey, 2000).⁶

Our interest in potential moderation (by deliberative features, gender, and NFC) created an analytic problem: Testing for all potential interactions would create type-1 error inflation and risk the identification of non-existent effects. On the other hand, neglecting to rule out interactions would not tell us whether the main effects (or lack of effects) observed in the data are robust across levels of other variables. We thus used the following analytic approach: We attempted to rule out interactions by starting with models including all 2- and 3-way interactions. We simplified these models using backward stepwise procedures to remove nonsignificant ($p > .05$) interactions one at a time beginning with 3-way, then 2-ways, and dropping those with largest p-values first. We did *not* retain interactions that were significant only when nonsignificant interactions were included in the model; however, if a higher-order interaction was retained, we also retained all related lower-level effects. If we were unable to drop all interactions using this procedure, then we do report the pattern of the significant interactions, but we also explicitly test for the interactions' existence across both studies to assess their reliability and potential importance.

Study 1

Engagement during reading: Enhanced by NFC but reduced by critical-thinking instructions. The preliminary MANCOVA analyses successfully ruled out all interactions predicting engagement during reading activities. The

⁵ The effect, found in S2 ($F(1,265) = 6.05, p = .015$), was that those in the critical-thinking condition scored lower on need for cognition ($M=4.26, SD=.67$) than those in the general engagement condition ($M=4.47, SD=.74$)

⁶ To facilitate readability, we use endnotes to report any preliminary analyses revealing faulty assumptions and how we dealt with those problems.

MANCOVA⁷ main effects model revealed a significant multivariate effect for reading-instruction condition (Wilk's lambda=.83, $F(12,304)=2.44$, $p=.005$, partial $\eta^2=.088$) and NFC (Wilk's lambda=.85, $F(6,152)=4.52$, $p<.001$, partial $\eta^2=.151$), but not for gender ($p=.497$). As shown in Table 1 (top half), univariate follow-ups revealed omnibus differences between reading conditions on all of the engagement scales except for the social engagement scale, with conscientiousness and boredom significant at Bonferroni-corrected levels. Contrary to our hypotheses, pairwise comparisons indicated the critical-thinking participants reported the *least* amounts of positive engagement and greatest amounts of negative engagement; and critical-thinking participants most often were significantly less engaged than those in the general engagement condition. Meanwhile, supporting the validity of the NFC and engagement measures, bivariate correlations between NFC and the engagement measures revealed positive relationships with active and conscientious engagement, and a negative relationship with closed-mindedness.

⁷ Box's M test for equality of variance-covariance was significant but not severely violated ($F(105,30465)=1.33$, $p=.013$).

Table 1

Study 1 univariate follow-ups examining the impacts of experimental conditions on engagement states during T2 reading and T3 deliberation activities

	NFC corr	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>p</i>	Partial eta ²
<i>T2 reading</i>										
		<u>Gen. eng. (n=51)</u>		<u>Info. org. (n=58)</u>		<u>Crit. Th. (n=53)</u>		<u>Df(2,157)</u>		
Conscientious	.25***	3.58 ^b	0.67	3.45 ^b	0.69	3.10 ^a	0.74	5.41	.005	.065***
Active/metacog.	.31***	3.21 ^b	0.79	2.87 ^a	0.82	2.75 ^a	0.95	3.43	.035	.042*
Social	.07	2.19	1.17	1.91	1.18	1.77	0.98	1.71	.185	.021
Boredom	-.03	2.49 ^b	0.85	2.56 ^a	0.98	3.21 ^b	1.17	8.12	.000	.094***
Angry	-.01	2.07 ^a	0.97	2.31 ^{ab}	1.03	2.66 ^b	1.10	3.99	.020	.048*
Closed-minded	-.22***	2.13 ^a	0.64	2.11 ^a	0.70	2.45 ^b	0.73	3.63	.029	.044*
<i>T3 deliberation</i>										
		<u>Gen. eng. (n=53)</u>		<u>Info. org. (n=60)</u>		<u>Crit. Th. (n=57)</u>		<u>Df(2,164)</u>		
Conscientious	.34***	3.88	0.71	3.67	0.67	3.53	0.61	2.98	.053	.035
Active/metacog.	.37***	3.50 ^b	0.64	3.08 ^a	0.67	3.07 ^a	0.85	5.44	.005	.062***
‡Social	.14	2.98	1.42	2.74	1.44	2.77	1.30	0.22	.803	.003
Boredom	-.21***	1.75	0.83	2.04	0.91	2.04	1.01	1.29	.278	.016
Angry	-.05	1.64 ^a	0.74	2.08 ^b	0.89	1.85 ^{ab}	0.88	3.33	.038	.039*
‡Closed-minded	-.07	2.08	0.58	2.27	0.65	2.22	0.68	1.19	.306	.014

(Table Continues)

Table 1 (continued)

	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>p</i>	Partial η^2
	<u>Individual (<i>n</i>=80)</u>		<u>Group (<i>n</i>=90)</u>		<u>Df(1,164)</u>		
Conscientious	3.54	0.71	3.82	0.62	6.75	.010	.040*
Active/metacog.	2.93	0.74	3.45	0.67	23.03	.000	.123***
‡Social	1.71	0.88	3.82	0.90	233.18	.000	.587***
Boredom	2.24	0.99	1.70	0.79	14.21	.000	.080***
Angry	2.08	0.87	1.68	0.80	8.42	.004	.049***
‡Closed-minded	2.33	0.70	2.08	0.57	5.93	.016	.035*

Notes. * $p < .05$, ** $p < .01$, *** $p < .008$ (cut-off for Bonferroni-corrected significance for six univariate follow-ups to the MANCOVA analysis). Means in the same row with common superscripts a or b are not significantly different at the uncorrected $p < .05$ level. Correlations with Need for Cognition (NFC) are simple (not partial) Pearson correlations.

‡See text and Figure 1 for more complex interactions involving social and closed-minded engagement.

Engagement during deliberation: Enhanced by discussion and NFC, but those who had been instructed to think critically or organize information lagged behind. When predicting engagement during deliberation, MANCOVA⁸ analyses revealed two interactions that we were unable to rule out: a social×gender×NFC interaction (Wilk's lambda=.90, $F(6,149)=2.70$, $p=.016$, partial $\eta^2=.098$), and a social×gender×instruction interaction (Wilk's lambda=.85, $F(12,298)=2.04$, $p=.021$, partial $\eta^2=.076$). Examination of the univariate followups indicated significant contributors to the first interaction were closed-minded ($p=.030$) and social engagement ($p=.015$), and the only significant contributor to the second interaction was closed-minded engagement ($p=.004$). To describe the other engagement states, we conducted a main effects MANCOVA excluding closed-minded and social engagement as dependent variables. That model revealed a significant main effects of reading-instruction condition (Wilk's lambda=.90, $F(8,322)=2.27$, $p=.023$, partial $\eta^2=.053$), social condition (Wilk's lambda=.83, $F(4,161)=8.30$, $p<.001$, partial $\eta^2=.171$), and NFC (Wilk's lambda=.84, $F(4,161)=7.82$, $p<.001$, partial $\eta^2=.163$), but not gender ($p=.301$).

Table 1 (lower half) reports means for each of the main effect conditions (social and closed-mindedness are still listed for completeness, although they were not in the MANCOVA model). Univariate analyses indicated active/metacognitive engagement varied during the deliberation, with those who had been in the critical-thinking and information-organization conditions during reading being significantly less engaged than those who had been in the general engagement condition. NFC again correlated with conscientious and active/metacognitive engagement, and this time also negatively correlated with boredom. Finally, comparison of social conditions found the discussion condition was generally more positively engaging and less negatively engaging than the individual condition.

Interaction follow-ups: Discussion increases social engagement and decreases closed-mindedness more consistently among women than men. We used multiple regression procedures to investigate the pattern of the interactions predicting closed-minded and social engagement, specifically examining the pattern of effects of the social manipulation on the engagement variables under low, medium, or high NFC and different instruction conditions, for men and women separately. As shown in Table 2, for *closed-minded engagement*, group discussion significantly suppressed closed-minded engagement among women low or medium in NFC, in both the general engagement and information organization conditions. For men, however, discussion more narrowly suppressed

⁸ Box's M test for equality of variance-covariance was significant but not severely violated ($F(231,13945)=1.17$, $p=.039$).

closed-mindedness only among those medium or high in NFC and in the critical-thinking condition. Meanwhile, the pattern of the social \times gender \times NFC interaction predicting *social engagement* was such that, for women, discussion had equally positive effects across levels of NFC (see Figure 1, white bars). For men, however, the positive impact of discussion on social engagement was less than for women on average and depended on NFC: As NFC increased, group discussion had a more positive impact on social engagement.

Table 2

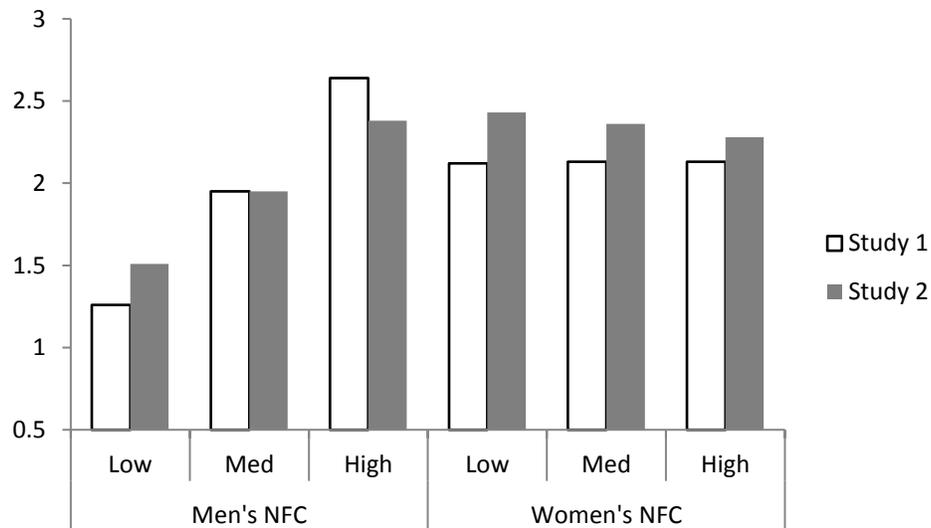
Study 1 gender-specific effects (unstandardized B weights) of discussion during deliberation activities on closed-minded engagement

	Men						Women					
	<u>Low NFC</u>		<u>Med NFC</u>		<u>High NFC</u>		<u>Low NFC</u>		<u>Med NFC</u>		<u>High NFC</u>	
	B	SE	B	SE	B	SE	B	SE	B	SE	B	SE
Gen. eng.	.59	(.30)	.33	(.25)	.06	(.27)	-.70*	(.28)	-.54*	(.25)	-.37	(.28)
Info. org	.11	(.27)	-.16	(.22)	-.43	(.25)	-.82**	(.29)	-.65*	(.25)	-.48	(.29)
Crit. th.	-.30	(.31)	-.57*	(.27)	-.84**	(.30)	-.01	(.24)	.16	(.21)	.32	(.27)

Notes. * $p < .05$, ** $p < .01$, *** $p < .001$. Gen. eng. = general-engagement condition, Info. org. = information-organization condition, Crit. th. = critical-thinking condition, NFC = need for cognition.

Figure 1

Impact (unstandardized beta weights) of discussion (vs. no discussion) on social engagement for men and women across levels of need for cognition (NFC)



Knowledge: Increases for all conditions about equally. Examination of pre-post knowledge change on repeated items indicated knowledge increased as expected, paired $t(166)=12.70$, $p<.001$, increasing from 45% ($SD=24\%$) to 70% ($SD=16\%$) correct. Preliminary MANCOVA⁹ analyses predicting both new and repeated post-knowledge scores from prior knowledge, NFC, gender, the effects of the reading and deliberation conditions, ruled out all interactions.¹⁰ The main

⁹ Preliminary analyses revealed two multivariate outliers based on Mahalanobis distance with $p<.001$ significance. We omitted the outliers because Box's M statistic was improved (to $p=.15$) with the two outliers excluded, and significance levels changed depending on inclusion (e.g., finding a $p<.05$ interaction with their inclusion, that was not significant ($p>.30$) when excluded).

¹⁰ After successfully ruling out other interactions using the backward stepwise procedures previously described, we also used multiple regression to explicitly test for the 3-way interactions that were observed when predicting social and closed-minded engagement. Only a marginal social \times gender \times instruction interaction was found ($F(2, 147)=2.60$, $p=.077$) when predicting proportion of new knowledge questions correct, with a pattern suggesting that discussion conditions marginally reduced post-knowledge among men in the information organization condition ($b=-.13$, $p=.085$). This may be useful to note primarily because it suggests less positive effects of discussion for men, consistent with other results.

effects model found that NFC positively predicted post-knowledge variables (Wilk's lambda=.886, $F(2,151)=9.68$, $p<.001$, partial $\eta^2=.114$), with follow-ups indicating it predicted both repeated and new knowledge scores (respective $F_s(1,152)=17.61$, 8.66 , $ps<.001$, $.004$, partial $\eta^2_s=.104$, $.054$). In addition, there was a marginal effect of prior knowledge (Wilk's lambda=.969, $F(2,151)=2.42$, $p=.093$, partial $\eta^2=.031$) with univariate follow-ups, indicating significance only when predicting the repeated knowledge questions $F(1,152)=4.53$, $p=.035$, partial $\eta^2=.029$). None of the experimental manipulations impacted post-knowledge after controlling for NFC and prior knowledge (all Wilk's $s>.975$, $ps>.15$).

Discussion of S1

The relationships between NFC and the engagement variables are consistent with the conceptualization of the conscientious and active/metacognitive measures as assessing deep and effortful cognitive engagement. However, our manipulations designed to increase deep engagement were ineffective. Asking participants to take organized pro-con notes had little to no impact on engagement or knowledge gains as measured in our studies, and our critical-thinking instructions appeared to disengage rather than engage participants. This 'disengagement effect' persisted after the reading task such that it was still detectable during the subsequent deliberation activities in reports of lower active/metacognitive engagement. For the reading conditions, the lack of interactions with NFC or gender suggest the deliberative reading activities were equally engaging across these individual differences.

Group discussion was found to be highly engaging, especially for women. Despite the generally positive effects of group discussion on self-reported engagement, overall, discussion did not impact knowledge as measured in this study. Also, while NFC did predict greater post-knowledge, neither NFC nor gender interacted with the experimental conditions to impact improved knowledge scores. Thus, the effects or lack of effects of the experimental conditions do not appear to vary based on gender or NFC.

Based on these results, as previously mentioned in the methods, we reduced the extensive critical-thinking instructions (which had included practice applying critical-thinking skills prior to reading) to instead consist of gentler prompts without practice. In addition, because our S1 information-organization condition did not significantly impact engagement or knowledge, and because practitioners likely have more control over information-presentation factors than how participants take notes, in S2 we made information-organization condition a presentation factor, randomly assigning participants to receive information pre-

organized in either a pro-con or topically organized format. Finally, we also expanded the types of engagement investigated, assessing creative and open-minded engagement in addition to the forms measured in S1.

Study 2

Engagement during reading: Enhanced by NFC and critical-thinking prompts; also was related to gender. Consistent with S1, no interactions were found predicting engagement while reading. Results from the main effects multivariate MANCOVA¹¹ indicated a significant multivariate effect for the instruction condition (Wilk's lambda=.75, $F(8,253)=10.34$, $p<.001$, partial $\eta^2=.246$), NFC (Wilk's lambda=.89, $F(8,253)=4.01$, $p<.001$, partial $\eta^2=.113$), and gender (Wilk's lambda=.93, $F(8,253)=2.39$, $p=.017$, partial $\eta^2=.070$), but not for information organization ($p=.457$). As shown in Table 3, females indicated less closed-minded and creative engagement than males. NFC again negatively correlated with boredom and positively correlated with active/metacognitive and conscientious engagement, as well as creative, open-mindedness and social engagement. *Positive* impacts of the critical-thinking condition on engagement were found for most of the dimensions as predicted. However, social engagement was slightly higher in the general-engagement than in the critical-thinking condition, and boredom was non-significantly higher in the critical-thinking condition.

¹¹ We identified six multivariate outliers based on Mahalanobis distance with $p<.001$ significance. Because inclusion of these outliers resulted in a severe Box's M statistic ($p<.001$ vs. $p=.04$ when deleted), we omitted them from the analyses.

Table 3

Study 2 univariate follow-ups examining engagement states during T2 reading

	Corr with Gender	Corr with NFC	General Engage. <i>n</i> =132		Critical Thinking <i>n</i> =133		Univariate Df(1,260)		Partial η^2
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>p</i>	
Conscientious	.00	.20***	3.55	0.71	3.92	0.61	27.42	.000	.095***
Open-minded	.04	.19***	3.77	0.76	4.05	0.66	14.90	.000	.054***
Active/metacog.	-.12	.21***	3.01	0.71	3.28	0.79	13.09	.000	.048***
Social	-.01	.15*	2.14	0.97	1.85	1.03	4.28	.040	.016*
Creative	-.13*	.25***	2.89	0.80	2.88	0.92	0.27	.606	.001
Boredom	.12	-.29***	2.16	0.74	2.34	0.78	1.62	.204	.006
Angry	-.01	-.05	1.57	0.68	1.42	0.70	3.83	.051	.015
Closed-minded	-.17***	-.03	2.22	0.81	1.85	0.87	15.04	.000	.055***

Notes. * $p < .05$, ** $p < .01$, *** $p < .006$ (cut off for Bonferroni-corrected significance for eight univariate follow-ups). For gender, 1=male, 2=female. Univariate tests are from the ANCOVA analysis, however descriptives are not adjusted for NFC and gender. Correlations are simple not partial correlations.

Engagement during deliberation: Enhanced by discussion, especially among women. Like S1, we were unable to rule out all interactions during S2. MANCOVA¹² analyses revealed a significant gender \times social interaction (Wilk's lambda=.92, $F(8,247)=2.67$, $p=.008$, partial $\eta^2=.080$), as well as main effects of gender (Wilk's lambda=.93, $F(8,247)=2.25$, $p=.024$, partial $\eta^2=.068$), NFC (Wilk's lambda=.92, $F(8,247)=2.74$, $p=.007$, partial $\eta^2=.081$), and discussion (Wilk's lambda=.34, $F(8,247)=60.66$, $p<.001$, partial $\eta^2=.663$). The information-organization ($p=.926$) and critical-thinking ($p=.793$) manipulations did not have significant effects.

As shown in Table 4, univariate follow-ups for individual engagement states indicated the gender \times social interaction involved conscientious, open-minded, and active/metacognitive engagement, which were promoted by discussion among women, but unaffected by social conditions among men. A similar but weaker pattern was observed for creative engagement. In addition, there was a tendency

¹² Eight multivariate outliers were identified examining Mahalanobis distances with $p<.001$ significance. Removal of the multivariate outliers and log transformations of angry (original skewness=3.03, kurtosis=10.78) and bored (original skewness=1.55, kurtosis=2.71) engagement to correct nonnormality resulted in a Box's M test that was acceptable ($F(540,21642)=1.17$, $p=.005$).

for discussion to increase closed-mindedness among men, but not women. For boredom, the interaction was not significant, but the univariate main effect of discussion was significant, $F(1,254)=5.00$, $p=.026$, partial $\eta^2=.019$), and the pattern of means resembled the interaction seen for other variables. For anger, the main effect of discussion was only marginal ($F(1,254)=3.52$, $p=.062$, partial $\eta^2=.014$), with discussion tending to reduce anger.

Although the multivariate tests did not reveal the 3-way interactions found in S1, given that those interactions only involved closed-minded and social engagement, they may have been hidden by the inclusion of other unaffected engagement states. Therefore, we explicitly tested for those interactions using univariate analyses. The interactions predicting closed-mindedness did not replicate. However, for social engagement, the univariate gender \times social \times NFC interaction was significant ($F(1,251)=5.44$, $p=.020$, partial $\eta^2=.021$). Follow-up analyses revealed a pattern similar to that found in S1: For men (but not women), as NFC increases, the positive effect of discussion increased (Figure 1, grey bars).

Knowledge: Generally increases during the engagement, with discussion (or social engagement) potentially benefitting women more than men. Mean correct repeated knowledge questions again significantly increased (from 58% (SD=19%) to 72% (SD=18%), $t(236)=9.64$, $p<.001$). MANCOVA¹³ analyses predicting scores on the repeated, new, and mixed knowledge questions allowed us to rule out all but one interaction (NFC \times prior knowledge, Wilk's lambda=.96, $F(3,225)=3.09$, $p=.028$, partial $\eta^2=.040$) and revealed overall main effects of NFC (Wilk's lambda=.96, $F(3,225)=3.45$, $p=.017$, partial $\eta^2=.044$) and prior knowledge (Wilk's lambda=.96, $F(3,225)=2.83$, $p=.040$, partial $\eta^2=.036$), but no effects of the instruction ($p=.716$), information ($p=.629$), or social ($p=.553$) conditions, or gender ($p=.324$).

¹³ For these analyses, Box's M was not significant ($p=.144$).

Table 4

Study 2 effects of NFC, group discussion, and gender on engagement states during T3 deliberation

Study 2	Corr with NFC	Individual (<i>n</i> =118)				Discussion (<i>n</i> =143)				Univariate		
		Male (<i>n</i> =54)		Female (<i>n</i> =64)		Male (<i>n</i> =56)		Female (<i>n</i> =87)		Group x Gender Df(1,254)		
		M	SD	M	SD	M	SD	M	SD	<i>F</i>	<i>p</i>	Partial eta ²
Conscien.	.17**	3.67 ^b	0.68	3.39 ^a	0.68	3.63 ^b	0.56	3.97 ^c	0.62	14.96	.000	.056***
Open	.17***	3.68 ^{ab}	0.72	3.48 ^a	0.71	3.80 ^b	0.69	4.11 ^c	0.60	8.76	.003	.033***
Active	.19***	3.10 ^b	0.65	2.70 ^a	0.74	3.26 ^{bc}	0.66	3.43 ^c	0.73	10.53	.001	.040***
‡Social	.01	1.85 ^a	0.98	1.75 ^a	0.86	3.85 ^b	0.86	4.11 ^b	0.64	3.02	.084	.012
Creative	.25***	3.18 ^b	0.73	2.74 ^a	0.73	3.37 ^b	0.70	3.38 ^b	0.79	5.98	.015	.023*
log Boredom	-.16*	0.21 ^{ab}	0.18	0.24 ^b	0.17	0.20 ^{ab}	0.15	0.15 ^a	0.16	3.21	.074	.012
log Angry	-.09	0.10 ^{ab}	0.13	0.12 ^b	0.16	0.08 ^{ab}	0.12	0.07 ^a	0.11	1.02	.313	.004
Closed	-.06	2.23 ^a	0.75	2.23 ^a	0.88	2.62 ^b	0.81	2.14 ^a	0.80	6.11	.014	.023*

Notes. * $p < .05$, ** $p < .01$, *** $p < .006$ (cut off for Bonferroni-corrected significance for eight univariate follow-ups). Means (uncorrected for NFC) in the same row with common superscripts are not significantly different at the uncorrected $p < .05$ level. Correlations are simple not partial correlations. Log transformations were applied to boredom and angry engagement to correct for nonnormality.

‡See text and Figure 1 for more complex interactions involving social engagement.

Examination of the univariate follow-ups suggested the NFC×prior knowledge interaction primarily involved the score from the new knowledge questions, $F(1,227)=3.77$, $p=.053$, partial $\eta^2=.016$. Regression results revealed a pattern such that prior knowledge was more predictive of new knowledge scores among those low in NFC ($b=.343$, $p=.004$, at 1 SD below the NFC mean) than among those high in NFC ($b=-.014$, $p=.912$, at 1 SD above the NFC mean).¹⁴ Examination of univariate main effects predicting mixed or repeated knowledge scores found that NFC positively predicted scores from mixed knowledge questions ($F(1,228)=12.42$, $p=.001$, partial $\eta^2=.052$) but not the repeated questions ($p=.157$), and prior knowledge predicted scores from the repeated questions ($F(1,228)=21.96$, $p<.001$, partial $\eta^2=.088$) but not the mixed questions ($p=.199$).

We conducted univariate analyses to again test for the 3-way effects found in S1 to predict closed-minded and social engagement (controlling for all experimental main effects, NFC, and prior knowledge). While those 3-ways were not significant, results did suggest the importance of the gender×social condition interaction when predicting scores from the repeated questions ($b=.101$, $SE=.046$, $t(227)=2.20$, $p=.029$). Participating in group discussion (compared to the individual condition) predicted lower scores for men ($b=-.077$, $SE=.035$, $t(227)=-2.21$, $p=.028$) but not for women ($b=.024$, $SE=.030$, $t(227)=8.12$, $p=.418$).

Finally, given the conceptual similarity between the discussion manipulation and self-reported social engagement, and in light of the prior findings of gender×social condition interaction predicting social engagement (Figure 1), in a separate analysis we tested for the interaction between gender, and self-reported social engagement predicting the knowledge variables (still controlling for all experimental main effects, NFC and prior knowledge). We found the interaction to be significant when predicting the scores of the repeated questions ($b=.052$, $SE=.023$, $t(221)=2.27$, $p=.024$) with the pattern of the interaction such that social engagement did not predict men's scores ($b=.002$, $SE=.022$, $t(221)=.094$, $p=.925$), but did predict higher scores for women ($b=.054$, $SE=.021$, $t(221)=-2.51$, $p=.013$).

General Discussion and Implications for Deliberative Practice

The present studies employed experimental methods to determine if features commonly viewed as essential to deliberative public engagements—balanced

¹⁴ This interaction did not replicate when explicitly tested in S1.

information presentation, deep cognitive engagement, and discussion—impacted participant cognitive-affective engagement and post-deliberation knowledge among college science students engaged in a deliberation about ethical issues related to nanotechnologies. We also examined the robustness of such effects across different levels of potentially important moderating factors; specifically, gender, and NFC. Our findings suggest three main points.

Deliberative Features Should be Tested in Different Contexts, Not Just Assumed

Features of deliberation theorized as essential may not be so in every context, at least for maximizing engagement and increasing knowledge. For example, we found little effect of information organization on cognitive-affective engagement or post-knowledge. This was true whether the participants (S1) or the background document (S2) did the organizing. Although our study occurred in an educational context (college students engaged in deliberation as part of their course), this finding goes against some prior educational research (e.g., as reviewed by Bohn-Gettler & Kendeou, 2014). Thus, it would be beneficial to conduct studies in other engagement contexts. For now, our results suggest a less-than-expected benefit from specific information presentation design such as advocated by groups like the National Issues Forum. It may be that information organization has other positive effects (such as greater awareness of a range of arguments¹⁵), but less impact on factual knowledge gains in deliberative contexts than in more traditional academic learning contexts because of the different goals. In American classrooms, students often expect to be tested over the information, whereas in deliberative contexts (including our classroom deliberation) knowledge is not commonly expected to be tested or in fact tested. Participants in deliberation do need to form and often explain and justify opinions; this may result in less or selective attention, possibly reducing overall impact of information or differences in text structures. Given that prior research finds complex interactions between individual differences, task goal, and text structure (Bohn-Gettler & Kendeou, 2014), additional research would be needed to identify conditions that might maximize the public’s learning during deliberation activities.

Attempts to Promote Deep Engagement Can Backfire

Efforts to enhance deep cognitive processing need to be designed carefully or may have opposite-of-intended effects. Our manipulation in S1 used rather intensive explanations and practice of critical-thinking skills during reading, and *negatively*

¹⁵ The authors thank an anonymous reviewer for this suggestion.

impacted self-reported engagement. One might have expected students in an academic context to have been more open to learning and practicing critical thinking given the academic motivations such a context promotes. However, the negative impacts of our instructions were detectable even during later discussion. This underscores the need for future research to test the impacts of such instructions in public engagement contexts. In S2, using simpler explanations and prompts to think carefully and critically, the critical-thinking manipulation had positive effects on engagement. Thus, there may be an optimal level at which to tune instructions so that they encourage effortful deliberation without undermining engagement, a finding which is consistent with other scholarship (Warnick et al., 2005). Importantly, the fine-tuning may depend on context and characteristics of the deliberative participants.

The Discussion Element of Deliberative Discussions May Impact Engagement More than Knowledge, and May Especially Benefit Women

Our studies found that discussion had large positive impacts on self-reported cognitive-affective engagement. However, discussion had very little effect on post-deliberation knowledge (with women benefitting more than men). Importantly, tests of our experimental manipulations on knowledge controlled for prior knowledge, which previously has been found to play a critical role in engagement with nanotechnology information-seeking and processing (Xenos, Becker, Anderson, Brossard, & Scheufele, 2011). Also, it is not likely that our lack of effects occurred due to invalid knowledge measures, as knowledge did increase overall across the activities and correlated with other variables, such as NFC and prior knowledge, as one might expect.

The lack of strong effects of discussion on knowledge is noteworthy, first, because it corroborates prior findings that most of the learning during deliberations comes from information provision, not discussion (Muhlberger & Weber, 2006); and second, because of the importance that some place upon public knowledge and understanding of science as an outcome (Powell & Kleinman, 2008). Our findings do not mean that all forms of discussion will be ineffective for all outcomes. Discussion may enhance forms of knowledge not assessed in this study or support better decision-making based on one's knowledge.¹⁶ However, our findings indicated that merely including discussion is not sufficient to ensure all positive outcomes. This is, of course, consistent with a great deal of other research, including Lewin's (1943) seminal research finding that group *decisions*, not just discussion, were key to changing behaviors.

¹⁶ The authors thank an anonymous reviewer for this suggestion.

Although there was little effect of discussion (or other experimental manipulations) on knowledge, there was some evidence that the effects of discussion were different for men versus women. Across both studies, interactions between gender and social conditions commonly emerged, with the pattern of the effects indicating that discussion was more consistently beneficial for women than men. In S1, discussion resulted in reports of reduced closed-minded engagement for women across a greater number of reading condition-by-NFC combinations. In both S1 and S2, discussion resulted in greater and more robust increases in social engagement for women, overall, than for men. Men also occasionally demonstrated less knowledge after discussion compared to individual reflection, whereas women did not. These findings are more consistent with Fraile's (2014) findings that deliberations may benefit women and reduce gender gaps, and less consistent with others' concerns about deliberation increasing gaps between under/over influential groups (Benhabib, 2002). While more research is needed to establish the conditions under which such effects are observed, it is possible that they occur because while "rational deliberation" is a stereotypical male activity, discussion (and consensus-building) is a stereotypical female activity. Supporting this interpretation, in S2 gender also interacted with self-reported social engagement such that, only for women, social engagement predicted higher knowledge assessed with the repeated knowledge questions.

Conclusions and Limitations

We recognize that although experimentation using undergraduates in a class context provides scientific power and control, it is not the same as policy deliberations that take place in the real world. Thus, an important limitation of these studies is the utilization of students as participants versus studying deliberations taking place in public contexts. Nevertheless, our controlled research shows the potential benefits and surprise findings that might emerge if systematic unpacking of public engagement features and processes were to be undertaken (PytlikZillig & Tomkins, 2011).

At a minimum, our studies demonstrate the value of incorporating experimental manipulations to identify features of deliberative engagements most important to specific outcomes, including outcomes related to public understanding of new scientific technologies. Some of the features of public engagements, although touted as essential, may not have strong impacts on learning or other desired outcomes. Thus, if science learning is an important objective of a deliberation, further research is needed to determine how best to maximize it.

Our studies also suggest the possibility that deliberations have more positive impacts, on engagement and learning, for women than men. This could be extremely important if confirmed in other research. Might there be cultural or other differences as well as gender effects?

In conclusion, these studies demonstrate there is a critical need and opportunity to undertake rigorous experimental research on the impacts of different public engagement design choices. Moreover, the studies suggest that there may be differences depending on specific outcomes that might be desired, and the research presented in this article indicates there may be critical moderators operating that could be important to understand and potentially to control.

References

- Barsoum, M. J., Sellers, P. J., Campbell, A. M., Heyer, L. J., & Paradise, C. J. (2013). Implementing recommendations for introductory biology by writing a new textbook. *CBE-Life Sciences Education*, *12*, 106-116.
- Benhabib, S. (2002). *The claims of culture: Equality and diversity in the global era*. Princeton, NJ: Princeton University Press.
- Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology*, *39*, 206-219.
- Burkhalter, S., Gastil, J., & Kelshaw, T. (2002). A conceptual definition and theoretical model of public deliberation in small face-to-face groups. *Communication Theory*, *12*, 398-422.
- Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Individual differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*, 197-253.
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*, 752-766.
- Chambers, S. (2003). Deliberative democratic theory. *Annual Review of Political Science*, *6*, 307-326.
- Chi, M. T., Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439-477.
- Chin, C., & Brown, D. E. (2000). Learning in science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching*, *37*, 109-138.
- Delli Carpini, M. X., Cook, F. L., & Jacobs, L. R. (2004). Public deliberation, discursive participation, and citizen engagement: A review of the empirical literature. *Annual Review of Political Science*, *7*, 315-344.

- Dinsmore, D. L., & Alexander, P. A. (2012). A critical discussion of deep and surface processing: What it means, how it is measured, the role of context, and model specification. *Educational Psychology Review*, 24, 499-567.
- Dow, B. J. J., & Wood, J. T. (2006). *The SAGE handbook of gender and communication*. Thousand Oaks, CA: Sage Publications.
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14, 4-58.
- Farrar, C., Fishkin, J., Green, D., List, C., Luskin, R., & Paluck, E. (2010). Disaggregating deliberation's effects: An experiment within a deliberative poll. *British Journal of Political Science*, 40, 333-347.
- Fishkin, J., & Luskin, R. (2005). Experimenting with a democratic Ideal: Deliberative Polling and public opinion. *Acta Politica*, 40, 284-298.
- Fraile, M. (2014). Does deliberation contribute to decreasing the gender gap in knowledge? *European Union Politics*, 1465116514527525.
- Fredricks, J. A., Blumenfeld, P., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59-109.
- Fulkerson, R. (1996). *Teaching the argument in writing*. Urbana, IL: National Council of Teachers of English.
- Garcia-Mila, M., Gilabert, S., Erduran, S., & Felton, M. (2013). The effect of argumentative task goal on the quality of argumentative discourse. *Science Education*, 97, 497-523.
- Goodin, R. E., & Niemeyer, S. J. (2003). When does deliberation begin? Internal reflection versus public discussion in deliberative democracy. *Political Studies*, 51, 627-649.
- Guston, D. H. (2014). Building the capacity for public engagement with science in the United States. *Public Understanding of Science*, 23, 53-59.

- Hickerson, A., & Gastil, J. (2008). Assessing the difference critique of deliberation: Gender, emotion, and the jury experience. *Communication Theory, 18*, 281-303.
- Huberty, C. J., & Petoskey, M. D. (2000). Multivariate analysis of variance and covariance. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 183-208). San Diego, CA: Academic Press.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Karpowitz, C. F., Mendelberg, T., & Shaker, L. (2012). Gender inequality in deliberative participation. *American Political Science Review, 106*, 533-547.
- Kauffman, D. F., Zhao, R., & Yang, Y.-S. (2011). Effects of online note taking formats and self-monitoring prompts on learning from online text: Using technology to enhance self-regulated learning. *Contemporary Educational Psychology, 36*, 313-322.
- Levin, B. B. (1995). Using the case method in teacher education: The role of discussion and experience in teachers' thinking about cases. *Teaching and Teacher Education, 11*(1), 63-79.
- Lewin, K. (1943). Forces behind food habits and methods of change. *Bulletin of the National Research Council, 108*, 35-65.
- Lukensmeyer, C. J., & Torres, L. H. (2006). *Public deliberation: A manager's guide to citizen engagement*. Washington, DC: IBM Center for the Business of Government.
- Luskin, R. C., Fishkin, J. S., & Jowell, R. (2002). Considered opinions: Deliberative polling in Britain. *British Journal of Political Science, 32*, 455-487.
- MacKuen, M., Wolak, J., Keele, L., & Marcus, G. E. (2010). Civic engagements: Resolute partisanship or reflective deliberation. *American Journal of Political Science, 54*, 440-458.
- Marcus, G. E., Neuman, W. R., & MacKuen, M. (2000). *Affective intelligence and political judgment*. Chicago: University of Chicago Press.

- Martin, G. P. (2012). Public deliberation in action: Emotion, inclusion and exclusion in participatory decision making. *Critical Social Policy*, 32(2), 163-183.
- McCoy, M. L., & Scully, P. L. (2002). Deliberative dialogue to expand civic engagement: What kind of talk does democracy need? *National Civic Review*, 91(2), 117-135.
- Mendelberg, T. (2002). The deliberative citizen: Theory and evidence. In M. X. Delli Carpini, L. Huddy, & R. Shapiro (Eds.), *Research in micropolitics: Political decisionmaking, deliberation, and participation* (Vol. 6, pp. 151-193). Greenwich, CT: JAI Press.
- Mercier, H., & Landemore, H. (2012). Reasoning is for arguing: Understanding the successes and failures of deliberation. *Political Psychology*, 33, 243-258.
- Morrell, M. E. (2005). Deliberation, democratic decision-making and internal political efficacy. *Political Behavior*, 27(1), 49-69.
- Muhlberger, P., & Weber, L. M. (2006). Lessons from the Virtual Agoral Project: The effects of agency, identity, information, and deliberation on political knowledge. *Journal of Public Deliberation*, 2(1), 1-39.
- Nussbaum, E. M., & Kardash, C. M. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology*, 97, 157-169.
- Powell, M., & Kleinman, D. L. (2008). Building citizen capacities for participation in nanotechnology decision-making: The democratic virtues of the Consensus Conference Model. *Public Understanding of Science*, 17(3), 329–348. doi: 10.1177/0963662506068000
- PytlikZillig, L. M., Hutchens, M., Muhlberger, P., Wang, S., Harris, R., Neiman, J., & Tomkins, A. J. (2013). The Varieties of Individual Engagement (VIE) measure: A confirmatory factor analysis across two samples and contexts. *Journal of Public Deliberation*, 9(2), Article 8. Retrieved from <http://www.publicdeliberation.net/jpd/vol9/iss2/art8>

- PytlikZillig, L. M., & Tomkins, A. J. (2011). Public engagement for informing science and technology policy: What do we know, what do we need to know, and how will we get there? *Review of Policy Research*, 28, 197-217.
- Robinson, D. H., & Kiewra, K. A. (1995). Visual argument: Graphic organizers are superior to outlines in improving learning from text. *Journal of Educational Psychology*, 87, 455-467.
- Ryfe, D. M. (2005). Does deliberative democracy work? *Annual Review of Political Science*, 8, 49-71.
- Sanders, D. (2012). The effects of deliberative polling in an EU-wide experiment: Five mechanisms in search of an explanation. *British Journal of Political Science*, 42, 617-640.
- Van Blankenstein, F. M., Dolmans, D. H., Van der Vleuten, C. P., & Schmidt, H. G. (2013). Relevant prior knowledge moderates the effect of elaboration during small group discussion on academic achievement. *Instructional Science*, 41, 729-744.
- Warnick, B., Xenos, M., Endres, D., & Gastil, J. (2005). Effects of campaign-to-user and text-based interactivity in political candidate campaign web sites. *Journal of Computer-Mediated Communication*, 10: 00.
doi: 10.1111/j.1083-6101.2005.tb00253.x
- Xenos, M. A., Becker, A. B., Anderson, A. A., Brossard, D., & Scheufele, D. A. (2011). Stimulating upstream engagement: An experimental study of nanotechnology information seeking. *Social Science Quarterly*, 92, 1191-1214.