

Central Lancashire Online Knowledge (CLoK)

Title	Contemplating the Opposition: Does a Personal Touch Matter?
Type	Article
URL	https://clock.uclan.ac.uk/id/eprint/30999/
DOI	https://doi.org/10.1080/0163853X.2019.1701918
Date	2020
Citation	Iordanou, Kalypso and Kuhn, Deanna (2020) Contemplating the Opposition: Does a Personal Touch Matter? Discourse Processes, 57 (4). pp. 343-359. ISSN 0163-853X
Creators	Iordanou, Kalypso and Kuhn, Deanna

It is advisable to refer to the publisher's version if you intend to cite from the work.
<https://doi.org/10.1080/0163853X.2019.1701918>

For information about Research at UCLan please go to <http://www.uclan.ac.uk/research/>

All outputs in CLoK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the <http://clock.uclan.ac.uk/policies/>

Article

Contemplating the Opposition: Does a Personal Touch Matter?

Iordanou, Kalypso and Kuhn, Deanna

Available at <http://clock.uclan.ac.uk/31003/>

*Iordanou, Kalypso ORCID: 0000-0001-5930-9393 and Kuhn, Deanna (2019)
Contemplating the Opposition: Does a Personal Touch Matter? Discourse Processes .*

It is advisable to refer to the publisher's version if you intend to cite from the work.

For more information about UCLan's research in this area go to <http://www.uclan.ac.uk/researchgroups/> and search for <name of research Group>.

For information about Research generally at UCLan please go to <http://www.uclan.ac.uk/research/>

All outputs in CLoK are protected by Intellectual Property Rights law, including Copyright law. Copyright, IPR and Moral Rights for the works on this site are retained by the individual authors and/or other copyright owners. Terms and conditions for use of this material are defined in the <http://clock.uclan.ac.uk/policies/>

Contemplating the Opposition: Does a Personal Touch Matter?

Cite:

Iordanou, K., & Kuhn, D. (in print). Contemplating the opposition: Does a personal touch matter? *Discourse Processes*.

Official URL: <https://doi.org/10.1080/0163853X.2019.1701918>

Contemplating the Opposition: Does a Personal Touch Matter?

Abstract

Is it important to hear positions opposing one's own from others who genuinely believe them, as claimed by J.S. Mill (1859)? We examine whether the thinking of those who engage in discourse with peers who hold an opposing view benefit by hearing arguments favoring the opposing position expressed by individuals known to hold this position. We report on 127 young adolescents who were given access to identical relevant evidence, and engaged in dialogs on gas vs solar energy, in preparation for a whole-class debate. In the (randomly assigned) experimental classroom electronic dialogs were conducted with a series of peers who held an opposing view; in the control classroom dialogs were confined to same-side peers. Differences in prevalence and types of functional evidence-based argumentative idea units in individual final essays on the topic favored the experimental group. Also, differences by condition in participants' choice of evidence to access during the preceding dialogs reflected differences in patterns of inquiry. Differences appeared as well in post-intervention essays on a non-discourse topic, suggesting the superior group had made gains in understanding argumentation itself. Extension of the study longitudinally to a second year with a new topic showed continued gains and condition differences, supporting this interpretation – with the experimental group surpassing the control group. Potential generalization to adults' discourse on topics involving higher affect and commitment is considered.

Contemplating the Opposition: Does a Personal Touch Matter?

Concern continues over increasing confinement of discourse to like-minded groups (Lukianoff & Haidt, 2018). Such discourse is not necessarily negative in its effects (Kuhn et al., 2019), yet inclusion of opposing views is desirable on multiple grounds, with potentially far-reaching effects on attitudes, cognition, and behavior (Hodson et al., 2018). Ideas expressed may be new to those hearing them and influence their thinking. Features of the delivery and the deliverers themselves may contribute to such effects. But does the mere existence and presence of a person holding an opposing view play a critical role? Will the receiver's thinking about the issue benefit more if new ideas are personally represented? Might it serve to emphasize that there indeed exists a flesh-and-blood other who espouses such views and that what is being encountered is the authentic view of such an individual? This hypothesis can be traced back to J. S. Mill (1859). Here we ask whether it is possible to put it to a rigorous empirical test.

Devising a method of doing so is not straightforward. One might compare effects of delivery of the same message in face-to-face discourse versus its delivery in a neutral form, removed from any interpersonal context, thereby eliminating influence of the individual delivering it. Yet there remain the attitudes and dispositions of the receiver, and in the case of political or social issues, these are likely to be substantial. They may create resistance to even engaging with the issue, especially if the deliverer is from an outgroup (Maier, Richter, Nauroth, & Gollwitzer, 2018). Indeed, such influences have figured prominently in the long history of research on belief change following purely informational interventions (see Baron, 1994, for review). We therefore sought to design an experimental comparison that minimized affective

CONTEMPLATING THE OPPOSITION

associations that both individuals bring to a dialog, as well as potential affective features of the experience they encounter there, without forsaking the discourse context itself.

We pursued these objectives by choosing discourse topics and a participant sample for whom such affective associations were likely to be similarly small, thereby minimizing them as potential contributors. In addition, participants conducted their discourse electronically, restricting influence of physical or other personal attributes of conversational partners. We do include two features essential to serious discourse – that it have a meaningful purpose and that participants engage deeply in it over a sustained period, in contrast to a superficial exchange in which they merely espouse their respective positions.

The topic was one in which new information would be needed in order for participants to engage in meaningful debate and make strong arguments for their position as superior to the alternative. Multiple pieces of such information were made available and remained so for participants to make use of as they chose. These were balanced with respect to potential to support or weaken either of the two positions. They were identical across two experimental conditions and presented in a neutral, non-interpersonal (written) form.

The experimental conditions differed only with respect to pairing of conversational partners. Those in the experimental (*Opposing*) condition engaged in a series of dialogs with different peers who held an opposing position to the one they favored. Those in a comparison (*Own*) condition engaged in a parallel series of dialogs with peers but confined to those who held the same position as their own. Participants in both conditions had equivalent opportunity to engage in discussion of the topic and to consider the new information and its implications as potential evidence in relation to either position. In both conditions, participants were told their

CONTEMPLATING THE OPPOSITION

dialogs were in preparation for a later full-group debate. Thus, both groups had an equivalent motive and purpose for discussing the topic.

The key research question the experimental design addresses is this: Does the thinking of those in the *Opposing* condition benefit by hearing arguments favoring the opposing position expressed by individuals known to hold this position? Alternatively, is the benefit equivalent among those in a comparison *Own* condition, who have access to the same evidence and thus potential arguments, but contemplate these only with their like-minded peers? The answer to this question has the potential to inform policy debates regarding the value of promoting dialog between groups holding opposing views, an implication we return to.

Following a review of the most directly relevant literature, we present an initial and follow-up study that address our key research question by means of the comparison of the experimental and control groups described above. The first study presents data of 131 middle-school students over the course of a year; the second study follows a reduced sample over a second year. We conclude with a general discussion that summarizes the main findings and considers their strengths, limitations, and implications.

Although our key research question does not derive directly from it, our study can be situated within the framework of the now large literature on argumentation (for reviews see Iordanou et al., 2016; O'Connor & Snow, 2018; Resnick et al., 2015, 2018). This literature can be divided in two broad categories, one addressing skills of argument and another addressing argument as a means of knowledge acquisition (Asterhan & Schwarz, 2016; Andriessen, Baker, & Suthers, 2003) – learning to argue vs. arguing to learn – although both of these objectives can be achievable within a single activity (Iordanou, et al., 2019). In the present work we thus assess the effect of our intervention with respect to both. We examine knowledge gains directly, as

CONTEMPLATING THE OPPOSITION

assessed in the individual final essays participants write on the topic (Rapanta et al., 2013) following the intervention, specifically the number and types of evidence-based claims the essay contains, using a previously established coding system for this purpose (Kuhn et al., 2016).

In addition we examine the effect of the intervention with regard to the argument skills reflected in these essays. Extended interventions of the form employed here have led to gains in argumentation discourse skill (Crowell & Kuhn, 2014; Iordanou, 2013; Iordanou & Constantinou, 2015; Kuhn, 2019) itself, as well as in its expression in individual written argument (Hemberger et al., 2017; Kuhn & Crowell, 2011; Shi, Matos & Kuhn, 2019), especially in the critical respect of addressing opposing arguments and seeking to weaken them, a key objective of skilled argumentation (Walton, 1989; Macagno, 2016). Thus, an additional hypothesis can be advanced here and the role of our condition manipulation examined in this regard. We test the hypothesis of skill gains by including in the design a transfer topic, in order to assess the extent to which any argument skill gains may have generalized beyond the particular topic addressed in the intervention.

In sum, we anticipate the present work to contribute to the now extensive literature on argumentation by adding clarification regarding the process by which discourse may enhance thinking about a particular topic and argumentative thinking more broadly. Is the informational component of discourse sufficient or does its human embodiment enhance its effect? To answer this question, two outcomes are examined. One is the extent to which the new informational content (knowledge) is recognized and employed in an argumentative function, in the form of evidence-based claims. The second, pertaining to argumentative skill, is the range of different types of argumentative functions employed, using the coding scheme noted above. Based on previous research using the dialogic intervention employed here (Iordanou et al., 2019;

CONTEMPLATING THE OPPOSITION

Hemberger et al., 2017; Kuhn & Crowell, 2011), we can predict skill as well as knowledge gains in both conditions. However, if skill gains in the experimental (*Opposing*) condition are no greater than those in the comparison (*Own*) condition, we can infer the intervention discourse to be serving a primarily informational role.

Method

Participants

Participants were 131 middle-school students (63 female) from a public school in a suburban area of a southern European city. This age group was chosen so as to minimize their having the long-standing or entrenched views more commonly found among older individuals. Participants were mostly from middle-class families and within an average range of ability and academic achievement. All participated as part of their science classes, taught by the same primary school teacher, who received training by the authors but was blind to the goals of the study. Four equivalent 5th grade and two equivalent 6th grade classes of 10- to 11-year olds, were each randomly assigned to the *Opposing* or *Own* conditions in a quasi-experimental design, yielding a total of 66 participants in the *Opposing* condition and 65 in the *Own* condition. Supporting the initial equivalence of groups was students' performance on the two initial essays, on the main topic of whether gas or solar energy should be used for electricity production and on the transfer topic of whether cigarette sales should be banned (see Preliminary Analysis section below). Due to absences, some students failed to write an essay for one of the topics at a particular time, with the result that group sizes varied by topic and time.

CONTEMPLATING THE OPPOSITION

Procedure

Initial assessment. Participants were asked to write a brief essay taking a position on whether their country should use gas or solar energy to produce electricity, a controversial national topic that students had some familiarity with. A brief description of each energy production method was provided. All students had received instruction regarding energy and alternative sources of producing electricity in the context of a unit in their school science curriculum. The prompt was to “Write the argument you would make to someone who didn't agree that your position is the better one.” For comparison purposes, participants also wrote an essay on a non-discourse topic, whether cigarette sales should be banned.

The intervention took place over twelve 90-min twice-weekly class periods that began the next week. All participants were told the purpose of the activity was to learn about the topic and to prepare for a whole-class verbal debate, which took place after the intervention and assessment was completed. Participants were assigned to one of two teams – gas or solar energy – based on the position they supported in the initial essay.

Phase 1: Preparation. In Sessions 1-3, participants in both conditions assembled, randomly, into same-side groups of 5-6 each, generated reasons supporting their side's preferred position and shared them with one another, reflected on their reasons, recording them on cards, consolidating and eliminating duplicates, and then ranked the reasons cards with respect to their strength.

Phase 2: Dialogs. In Sessions 4-11, each team – gas or solar energy – divided into same-side pairs who remained together throughout the next sessions. Beginning with Session 4, pairs were presented with an information bank in the form of a set of cards, each of which contained a question on one side and a short answer to the question concealed on the other side. At the

CONTEMPLATING THE OPPOSITION

beginning of each of the eight sessions, before the dialog began, the same-side pair was able to choose three cards and access the answers. At each new session, the cards they had chosen in previous sessions remained available to them and three new ones could be chosen. By the end of these sessions, all participants had seen an identical set of 24 Q&A cards. The set was balanced overall with respect to support of the two positions. One, for example, asked whether it is expensive to produce electricity by solar power and another what the cost is of producing electricity using natural gas. Some pieces were more explicit than others in their potential function in an argument; two, for example, were “What is the most desirable feature of natural gas as energy?” and “What kinds of bad effects can natural gas have?”

During each of Sessions 4-11, a same-side pair engaged in an electronic dialog via an instant-messaging platform with a different pair from their classroom. In the *Opposing* condition, the series of pairs, one per session, were always pairs from the opposing side. In the *Own* condition, the pairs were always ones from the same side. Participants were instructed to introduce themselves at the beginning of the dialog and collaborate with their partner to decide what they wished to communicate electronically to the other pair they were to engage in dialog with. Participants received no explicit instruction with respect to argumentation.

While waiting for the other pair to respond, pairs reflected on an electronic transcript of the dialog and completed one of two reflection sheets, alternated across sessions. Own-side reflection sheets asked participants to reflect on the opposing side’s position and on the strength of the counterarguments they had constructed to the opposing side’s position. Other-side reflection sheets asked participants to reflect on the rebuttals they used to weaken counterarguments to their own position. In each case, they were asked to contemplate what a better counterargument or rebuttal might have been. Thus, participants in both conditions had

CONTEMPLATING THE OPPOSITION

equivalent opportunities to construct arguments, counterarguments, and rebuttals for both positions.

Participants in the *Own* condition, who conducted their electronic dialogs only with same-side pairs, were instructed to discuss with the pair they were matched with at that session to a) examine how the new information now available could be used, b) together construct their arguments, c) anticipate what the opposing side will say, and d) plan how to respond to it.

Participants in the *Opposing* condition similarly examined the new information and then engaged in electronic dialogs with other-side pairs. They were instructed to try to convince the other pair that their position was the better one.

Phase 3: Review. In Session 12, same-side groups reassembled and worked together to prepare for the ensuing whole-class debate. The group reviewed and reflected on the reflection sheets they had completed in previous sessions to decide what arguments to use in the debate. An adult coach (a research assistant) facilitated these discussions. Students were encouraged to use different colored cards to summarize arguments, counterarguments and rebuttals for both sides and to include evidence to support each.

Final assessment. One week following Session 12, all participants were asked to write a final essay on the discourse topic. The prompt again was to “Write the argument you would make to someone who didn't agree that your position is the better one.” Participants also again wrote an essay on a non-discourse topic (banning cigarette sales).

Results

An examination of the idea units in initial and final individual essays served as the main indicator in assessing students' achievement and in comparison of differences across conditions.

Segmentation and coding of essays

CONTEMPLATING THE OPPOSITION

Each initial and final essay was divided into idea units, with an idea unit defined as an assertion (e.g., “I am against natural gas”) with accompanying justification (e.g., “natural gas is expensive”). One of the authors and another coder, blind to condition and time, segmented into idea units the 508 initial and final essays participants wrote on the discourse and transfer topics. Inter-rater reliability on segmenting was achieved on a subset of 25% of units with 93% agreement, and the first author proceeded with segmenting the remaining essays, again blind to condition and time. Only segments that included a claim and supporting evidence (evidence-based units) were included in the data base for further analyses. A statement was coded as evidence if it was at least potentially empirically verifiable and offered an explicit or implicit answer to the question “How do you know?”. The majority of the evidence came from the Q&A items provided to participants, but some participants also included other evidence known to them. Occasionally participants cited a piece of information from one of the Q&A items provided, more or less verbatim, without connecting it to any claim. These were coded as non-functional units and not analyzed further. Occasionally a connection appeared to be attempted but was not clear enough to be discernible, in which case the unit was also classified as non-functional and not included in the data analysis.

Remaining units, which contained a claim and accompanying evidence used to support (or weaken) it, were classified as functional units. These were further classified as serving one of four functions, using a coding scheme adapted from that reported by Kuhn and colleagues (Kuhn, Hemberger, & Khait, 2016; Hemberger, Kuhn, Matos, & Shi, 2017), based on the rationale that skilled argument requires attention to all four argument functions and the coordination of units serving contrasting functions. Table 1 shows the four types of possible functions that functional units might serve, along with examples of each.

Table 1

Inter-rater reliability for identification and coding of functional units classified as serving one of the four functions was achieved on a subset of 25% of units with Cohen's kappa .91 ($p < .001$); differences were resolved by discussion and one of the authors proceeded with segmenting and coding the remaining essays, again blind to condition and time.

Preliminary analysis

Data for both the discourse topic and the non-discourse topic were first screened for normality. All variables – mean number of idea units, functional evidence-based claims, and the four subtypes (Support-own, Weaken-other, Support-other and Weaken-own) – followed the Poisson distribution. Mann-Whitney U test on participants' initial essays on both the energy topic and the cigarette topic showed no significant difference between *Opposing* and *Own* conditions at initial assessment in the number of idea units ($U = 1997.500, p = .928$, and $U = 1901.000, p = .247$), functional usage ($U = 1826.00, p = .341$, and $U = 1990.50, p = .555$), Support-own usage ($U = 1840.50; p = .376$, and $U = 1887.00, p = .266$), Weaken-other usage ($U = 1981.00, p = .669$ and $U = 2100.50, p = .934$), Support-other usage ($U = 1953.00, p = .159$, and $U = 2112.00, p = .991$), and Weaken-own usage ($U = 1984.00, p = .313$, and $U = 2112.50, p = 1.00$), respectively.

A generalized linear mixed model (GLMM) using the Poisson probability distribution overall test for number of idea units, including both functional and non-functional ones, was employed for the purpose of assessing effects on total relevant essay output. The overall model was significant, $F(3, 256) = 35.085, p < .001$. The interaction between time and condition was not significant, $F(1, 256) = .343, p = .559$, nor was the fixed effect of condition, $F(1, 256) = 0.767, p = .382$. The fixed effect of time (initial vs. final essay) was significant, $F(1, 256) = 104.394, p < .001$. The mean number of idea units in participants' essays increased from 2.039

CONTEMPLATING THE OPPOSITION

($SD = 1.723$) at initial assessment to 4.731 ($SD = 2.455$) at the final assessment (95% CIs 1.739, 2.337 and 4.305, 5.157), respectively.

Functional evidence-based idea units by time and condition

Discourse topic. How successful were participants in producing functional evidence-based ideas in their individual final essays on the discourse topic? In the *Opposing* condition, the percentage of participants who produced *any* functional evidence-based idea units increased from 68% in the initial essay to 94% in the final essay, a significant increase (McNemar, $p < .001$). In the *Own* condition, this percentage similarly was 63%, in the initial essay but remained at nearly the same level, 62%, in the final essay (see Table 2). (A Generalized Estimating Equations (GEE) with binary responses showed that the interaction between condition and time was not significant, $p = .052$.) Most common among unsuccessful participants in the final essay was simply to insert a direct quote of one of the evidence pieces, without attributing any meaning to it by connecting it to a claim.

Also of interest is the number of such functional evidence-based claims essays contained. Despite the superior posttest performance of the *Opposing* group, a generalized linear mixed model (GLMM) using the Poisson probability distribution showed that the overall model was significant, $F(3, 256) = 23.557$, $p < .001$, but that the interaction between time and condition was not significant, $F(1, 256) = 1.571$, $p = .211$. Nor was the main effect of condition, $F(1, 256) = 1.163$, $p = .282$. Overall effect of time was significant, however, $F(1, 256) = 69.974$, $p < .001$, with mean number of functional units increasing from 1.423 ($SD = 1.642$), 95% CI [1.138, 1.708], at initial assessment to 3.631 ($SD = 2.500$), 95% CI [3.197, 4.065] at final assessment.

CONTEMPLATING THE OPPOSITION

Thus, once *Own* condition participants achieved success in making functional evidence-based claims, they were no less successful than *Opposing* condition participants in producing them.

Non-discourse topic. A similar analysis was conducted for the non-discourse topic. The percentage of participants in the *Opposing* condition who produced *any* functional evidence-based idea units increased significantly from 81% in the initial essay to 97%, in the final essay (see Table 3). In the *Own* condition, this percentage also increased from 70%, in the initial essay to 86%, in the final essay. (The slightly higher initial percentages, compared to the discourse topic, are likely attributable to the fact that the smoking topic was one for which participants had somewhat more personal knowledge to bring to bear.) (A GEE analysis with binary responses showed that the interaction between condition and time was not significant, $p = .280$.)

A GLMM using the Poisson probability distribution showed a difference between conditions over time in mean frequencies of functional usage, $F(3, 246) = 33.682, p < .001$. The interaction between condition and time was statistically significant $F(1, 246) = 11.761, p < .001$. The fixed effect of condition was significant in the overall model, $F(1, 246) = 18.233, p < .001$, as was the fixed effect of time, $F(1, 246) = 93.545, p < .001$. Participants in the two conditions showed similar initial means – 1.540 (1.255 s.d.) for the *Opposing* condition and 1.359 (1.302 s.d) for the *Own* condition, 95% CIs [1.224, 1.856] and [1.034, 1.685], respectively; for the final essay these diverged, with superior performance by the *Opposing* condition in essays on the topic that had not been part of the intervention – 4.350 (1.990) for the *Opposing* condition and 2.719 (1.915) for the *Own* condition, 95% CIs [3.836, 4.864] and [2.214, 3.183], respectively.

Types of functional evidence-based idea units

CONTEMPLATING THE OPPOSITION

What functions did these evidence-based claims serve? Unsurprisingly, and consistently across individuals, times, and condition, the majority of idea units in essays employed evidence to serve the purpose of supporting a claim consistent with the writer's position on the topic. Given an objective of discourse is to consider another's ideas, of greatest interest are idea units that had functions other than to support one's own position. These other functions could be ones that address opposing ideas, either critically (Weaken-other) or acknowledging and addressing their merit (Support-other). They could also be ones that address one's own position critically by addressing its weaknesses (Weaken-own).

Weaken-other functions. Of the three types just indicated, we consider Weaken-other idea units first as they were the most prevalent of the three – unsurprising since, unlike the other two, these do not threaten one's own position. Unlike Support-own units, however, which appeared in nearly all final essays in the *Opposing* condition, Weaken-other units did not reach this level of universality in either condition, a fact consistent with the challenge young and even older writers are known to have in writing two-sided essays (Newell et al., 2011).

For the discourse topic, appearance of *any* Weaken-other units began at an equally low level (19% of participants in each condition) and rose to only 62% of the *Opposing* and 63% of the *Own* group including even one such unit in their final essays. (A GEE with binary responses showed that the interaction between condition and time was not significant, $p = .740$.)

A GLMM using the Poisson probability distribution for number of such units showed that the overall model was significant in the discourse topic, $F(3, 256) = 24.285, p < .001$. The interaction of time and condition was not significant, $F(1, 256) = 1.610, p = .206$; neither was the fixed effect of condition, $F(1, 256) = 0.754, p = .386$. The fixed effect of time was significant, $F(1, 256) = 71.949, p < .001$. Mean number of units increased from 0.208 ($SD = 0.477$), 95% CI

CONTEMPLATING THE OPPOSITION

[0.125, 0.291] at initial assessment, to 1.185 ($SD = 1.231$), 95% CI [0.971, 1.398] at final assessment.

For the non-discourse topic, participants were even less likely to include Weaken-other units in their essays. Only three participants (5%) in the *Own* condition did so at either time. The *Opposing* group, showed an increase from four (6%) doing so at the initial time to 11 (18%) at the final assessment. (A GEE analysis with binary responses showed that the interaction between condition and time was not significant, $p = .329$). A GLMM using the Poisson probability distribution for number of Weaken-other units in the non-discourse topic showed that the overall model was not significant, $F(3, 246) = 1.267$, $p = .286$, neither were the interaction term, $F(1, 246) = 0.402$, $p = .527$, nor the fixed effect of condition, $F(1, 246) = 1.919$, $p = .167$, or time, $F(1, 246) = 1.919$, $p = .167$. The Weaken-other usage was not frequently used by *Opposing* and *Own* condition participants at either the initial assessment ($M = 0.063$, $SD = 0.275$, 95% CI = 0.015, 0.111) or at the final assessment ($M = 0.122$, $SD = 0.396$, 95% CI = 0.051, 0.193).

In sum, both groups were challenged to produce Weaken-other units and only about two thirds succeeded, even following extended discourse on the topic. Success on a non-discourse topic was much less prevalent, achieved by only a small minority, yet it was more prevalent among participants in the *Opposing* condition.

Support-other and Weaken-own functions. Although acknowledging and addressing them is essential to skilled argument, units serving the Support-other and Weaken-own functions were rarest of all, unsurprising given these have the potential to challenge and possibly weaken one's own position, rather than advance it. When they did occur, Support-other and Weaken-own statements were typically followed by a counterargument seeking to weaken their force. If we consider both types of position-challenging units (Weaken-own and Support-other) in total, for

CONTEMPLATING THE OPPOSITION

the discourse topic the number of participants ever producing either remained negligible initially but rose to 16 in the final essay in the *Opposing* condition but reached only four in the *Own* condition, a significant increase only for the *Opposing* condition ($p < .001$). In the transfer essay, the numbers remained equally low across conditions initially and flat across time.

Weaken-own functions. For the discourse topic, a Weaken-own unit appeared in the initial essay of one participant in the *Opposing* condition but rose to 10 participants in final essays. In the *Own* condition, both incidence and change remained minimal – from one to two. (A GEE analysis with binary responses showed that the interaction between condition and time was not significant, $p = .269$.) A GLMM using the Poisson probability distribution did show a difference between conditions over time in mean frequencies of units serving the Weaken-own function in the discourse topic, $F(3, 256) = 4.673$, $p = .003$. The interaction between condition and time was statistically significant $F(1, 256) = 5.095$, $p = .025$. The fixed effect of condition was significant in the overall model, $F(1, 256) = 5.095$, $p = .025$, as was the fixed effect of time, $F(1, 256) = 7.960$, $p = .005$. Pairwise contrasts showed that only students in the *Opposing* condition showed an increase in the Weaken-own usage from initial assessment ($M = 0.015$, $SD = 0.124$) to final assessment ($M = 0.154$, $SD = 0.364$), 95% CIs $[-0.015, 0.046]$ and $[0.064, 0.244]$, respectively. No significant change was observed from initial assessment ($M = 0.015$, $SD = 0.124$) to final assessment ($M = 0.031$, $SD = 0.174$) in usage of the Weaken-own type in the *Own* condition, 95% CIs $[-0.015, 0.046]$ and $[-0.012, 0.074]$, respectively.

For the non-discourse topic, none of the *Opposing* condition students used Weaken-Own units at initial assessment and only one did so at the final assessment. Similarly, none of the *Own* condition students used Weaken-One units at either the initial or the final assessment. A GLMM using the Poisson probability distribution in mean frequencies of Weaken-own units, in the non-

CONTEMPLATING THE OPPOSITION

discourse topic, showed that the overall model was not significant, $F(3, 246) = 0.689, p = .560$, neither were the interaction term, $F(1, 246) = 0.001, p = .972$, nor the fixed term of condition, $F(1, 246) = 2.066, p = .152$ nor the fixed term of time, $F(1, 246) = 0.001, p = .972$. Participants in both conditions rarely used Weaken-own units at either initial ($M = 0.007, SD = 0.089, 95\% CI = -0.008, 0.024$) or final assessment ($M = 0.008, SD = 0.090, 95\% CI = -0.008, 0.024$).

Support-other functions. For the discourse topic, a Support-other unit appeared in the initial essay of only one participant in each condition and in the final essays rose to six participants in the *Opposing* condition and two in the *Own* condition. (A GEE analysis with binary responses showed that the interaction between condition and time was not significant, $p = .455$.)

A GLMM using the Poisson probability distribution for the mean frequencies of Support-other units in the discourse topic showed that the overall model was not significant, $F(3, 256) = 2.068, p = .105$. Neither the interaction, $F(1, 256) = 2.064, p = .152$, nor the main effects of condition, $F(1, 256) = 3.083, p = .08$, and time, $F(1, 256) = 3.083, p = .08$, were significant. Participants in the *Opposing* and *Own* conditions showed limited usage of the Support-other type at either initial assessment ($M = 0.031, SD = 0.248, 95\% CI = -0.031, 0.092; M = 0.015, SD = 0.124, 95\% CI = -0.015, 0.046$) or final assessment ($M = 0.185, SD = 0.705, 95\% CI = 0.010, 0.359; M = 0.031, SD = 0.174, 95\% CI = -0.012, 0.074$, respectively).

For the non-discourse topic, none of the *Opposing* condition participants used Support Other units at initial assessment and only one did so at the final assessment. The frequencies were also low for the *Own* condition participants – only 2 used support other units at both initial and final assessments. A GLMM using the Poisson distribution for Support-other units in the non-discourse topic showed the model was not significant $F(3, 246) = 1.312, p = .271$; neither

CONTEMPLATING THE OPPOSITION

were the interaction term, $F(1, 246) = 0.000$, $p = .988$, nor the fixed effect of condition $F(1, 246) = 3.933$, $p = .05$, nor the fixed effect of time, $F(1, 246) = 0.000$, $p = .988$. Participants rarely used Support-other units at either initial ($M = 0.016$, $SD = 0.125$, 95% CI = -0.006, 0.038) or final assessments ($M = 0.016$, $SD = 0.127$, 95% CI = -0.006, 0.039).

Table 2

Table 3

Qualitative data on information choice

Pairs of participants in both conditions were asked to select three questions, of the set made available at the beginning of each dialog session, that they could then access answers to. Qualitative data on differences across conditions in choice behavior proved revealing. Six pairs (12 participants) in each condition were randomly selected for identification by an observer of the questions the pair chose (the maximum number of pairs the observer could document in the time available). Half of the questions concerned the solar alternative and half the fossil fuel alternative. A summary of these pairs' choice appears in Table 4 by condition and separately for the early (first four) and later (last four) sessions.

As seen in Table 4, patterns of information choice differed across conditions. The *Opposing* condition participants tended to initially explore the opposing alternative; they then in the later sessions turned to a focus on their own position. The *Own* condition participants in contrast initially preferred to seek information related to their own position and only later did some show a shift toward concentrating their information seeking on the opposing position.

Table 4

Discussion

Given the results of Study 1 supported both hypotheses, showing both pre-post gains and condition differences, as well as some generalization to a non-discourse topic, we decided to follow students for a second year to determine if these trends were maintained. We postpone more general discussion until Study 2 has been presented.

Study 2

Method

Participants

A total of 44 students who had participated in Study 1 were available to participate in Study 2. They were from the two fifth-grade classes (21 from one class and 23 from the other) who continued at the school as sixth graders. Participants remained in the same *Opposing* (n = 21) or *Own* (n = 23) conditions they had served in during their first year of participation.

Procedure

The procedure in Year 2 was identical to that of Year 1. The discourse topic used in the second year was whether Genetically Modified Food (GMF) should be allowed or not. The non-discourse topic was the same cigarette topic used in Year 1. Pre- and posttest assessments were similarly conducted prior to and following the intervention.

Results

Measures for this sample are examined across three time points over two years: at the beginning of the first intervention in year 1 (T1), at the end of the first intervention and before

CONTEMPLATING THE OPPOSITION

the beginning of the second intervention (T2), and at the end of the second intervention in year 2 (T3). A generalized linear mixed model (GLMM) using the Poisson probability distribution was again used to compare the two conditions. (Omitted are evidence-selection performance data, which it was not possible for the teacher to record in the second year.)

Functional evidence-based idea units by time and condition

The gains observed at the end of Year 1 regarding the functional usage of evidence remained evident in Year 2. All participants in the *Opposing* condition and almost all participants in the *Own* condition, 96%, produced at least one functional evidence-based idea unit in the discourse topic. Similarly, 95% of the participants in the *Opposing* condition and 86% of the participants in the *Own* condition produced functional evidence-based idea units in the non-discourse topic.

Discourse topic. A GLMM using the Poisson probability distribution in functional units for the discourse topic showed that the overall model was significant, $F(5, 123) = 24.005, p < .001$. The interaction between time and condition was not significant, $F(2, 123) = 2.011, p = .138$, neither was the fixed effect of condition, $F(1, 123) = 2.250, p = .136$. Yet, the fixed effect of time was significant, $F(2, 123) = 59.620, p < .001$. Pairwise contrasts showed that the mean number of functional evidence-based units produced at T2 ($M = 3.864, SD = 2.611, 95\% CI = 3.069, 4.658$) was greater than at T1 ($M = 0.705, SD = 1.091, 95\% CI = 0.373, 1.036$). No significant difference was observed between T2 and T3, ($M = 4.756, SD = 2.773, 95\% CI = 3.880, 5.631$). The mean number of functional evidence-based units produced at T3 was greater than at T1.

CONTEMPLATING THE OPPOSITION

Non-discourse topic. A GLMM using the Poisson probability distribution overall test revealed a difference between conditions in functional units $F(5, 118) = 10.756, p < .001$. The interaction between condition and time was significant, $F(2, 118) = 3.545, p = .032$. The fixed effect of time was significant, $F(2, 118) = 24.636, p < .001$, while the fixed effect of condition was not significant, $F(1, 118) = 2.123, p = .148$. Pairwise contrasts showed a mean difference for both conditions between T2 and T1 and between T3 and T1 (see Figure 1); no significant difference was observed between T2 and T3. In particular, those in the *Opposing* condition produced more functional units at T2 ($M = 3.700, SD = 1.261, 95\% CI = 3.110, 4.290$) compared with T1 ($M = 1.400, SD = 1.429, 95\% CI = 0.731, 2.069$) and in T3 ($M = 4.211, SD = 1.960, 95\% CI = 3.266, 5.155$) compared again with T1. Similarly, those in the *Own* condition produced more functional units at T2 ($M = 3.591, SD = 2.260, 95\% CI = 2.589, 4.593$) and T3 ($M = 2.667, SD = 1.683, 95\% CI = 1.901, 3.433$) in comparison with T1 ($M = 1.727, SD = 1.032, 95\% CI = 1.269, 2.184$). At the end of Year 2 (T3) the *Opposing* condition produced more functional units ($M = 4.211, SD = 1.960, 95\% CI = 3.266, 5.155$) than the *Own* condition ($M = 2.667, SD = 1.683, 95\% CI = 1.901, 3.433$). In sum, for both discourse and non-discourse topics, students maintained their gains working with a new topic in the second year.

Types of functional evidence-based idea units

Weaken-other usage. For the discourse topic, 79% of *Opposing* condition participants and 52% of *Own* condition participants produced Weaken-other units. These percentages were comparable to the ones observed at the end of Year 1 (62% and 63% respectively). Although no significant increase was observed from Year 1 to Year 2 (T2 vs T3), there was a significant increase over time (from T1 to T3), ($p < .001$, Cochran's Q), for both conditions. A GLMM

CONTEMPLATING THE OPPOSITION

using the Poisson distribution showed the model for Weaken-other usage was significant, $F(5, 124) = 11.855, p < .001$. The interaction between time and condition was not significant, $F(2, 124) = .656, p = .520$, neither was the fixed effect of condition, $F(1, 124) = 1.541, p = .217$. The fixed effect of time was significant, $F(2, 124) = 29.547, p < .001$. Pairwise contrasts showed that the mean number of units produced at T2 ($M = 1.296, SD = 1.212, 95\% CI = 0.927, 1.664$) was greater than at T1 ($M = 0.136, SD = 0.347, 95\% CI = 0.031, 0.242$). Also, the mean number of units produced at T3 ($M = 1.191, SD = 1.215, 95\% CI = 0.812, 1.569$) was greater than at T1. No significant difference was observed between T2 and T3.

For the non-discourse topic, as was the case in Year 1 participants were less likely to include Weaken-other units in their essays. Yet the difference observed between conditions in Year 1 was more pronounced in Year 2. Although a small percentage (5%) of *Own* condition students produced Weaken-other units in Year 2 (the same percentage observed at the end of Year 1), almost a third of *Opposing* condition students – 32% – did so. The percentage of *Opposing* condition students who included Weaken-other units increased from Year 1 to Year 2 – from 10% to 32% – showing that the gains continued to accrue over time in the *Opposing* condition ($p = .009$, Cochran's Q). A GLMM using the Poisson distribution showed a difference between conditions over time in mean frequencies of Weaken-other usage, $F(5, 118) = 2.727, p = .023$. An interaction occurred between condition and time, $F(2, 118) = 3.533, p = .032$. Both the fixed effect of time and the fixed effect of condition were significant in the overall model: fixed effect of time, $F(2, 118) = 3.532, p = .032$; fixed effect of condition, $F(1, 118) = 5.019, p = .027$. Pairwise contrasts revealed a mean difference between T3 and T2 ($M = 0.100, SD = 0.308, 95\% CI = -0.044, 0.244$) as well as between T3 and T1 for the *Opposing* condition only. In sum, as reported above for total functional evidenced-based units, for both discourse and non-

CONTEMPLATING THE OPPOSITION

discourse topics students maintained their gains working with a new topic in the second year. Furthermore, condition differences became more pronounced.

Weaken-own usage. Weaken-own and Support-other units were the least frequent units in Year 1, with no more than 20% of participants showing them. Notably, at the end of Year 2, half of the *Opposing* condition participants, 53%, included units in their essays on the discourse topic that functioned as weakening their position, compared to less than 10% in the *Own* condition. *Opposing* condition students showed a significant increase over time, from T1 to T3, in the percentage of students using Weaken-own units (Cochran's Q, $p = .001$), as well as from Year 1 to Year 2 (T2 vs T3), (McNemar, $p = .031$). A GLMM using the Poisson probability distribution showed that the overall model was significant, $F(4, 104) = 5.632, p < .001$, revealing condition differences. Also the interaction between time and condition was significant, $F(1, 104) = 7.402, p = .008$. The fixed effect of time was significant, $F(2, 104) = 32.529, p < .001$. The fixed effect of condition was not significant, $F(1, 104) = 3.921, p = .050$. Pairwise contrasts showed significant improvement over time in the *Opposing* condition, but not in the *Own* condition. In particular, mean frequency of the Weaken-own strategy in the *Opposing* condition was greater at T3 ($M = 1.000, SD = 1.247, 95\% CI = 0.399, 1.601$) compared to both T2 ($M = 0.191, SD = 0.402, 95\% CI = 0.007, 0.374$) and T1 ($M = 0$). The results of the GLMM show that not only a greater percentage of *Opposing* condition students employed the Weaken-own strategy compared to *Own* condition in Year 2, but also that *Opposing* condition students employed this strategy more frequently over time and in comparison to the *Own* condition students.

CONTEMPLATING THE OPPOSITION

In the non-discourse topic, only a small proportion of students (11% and 5% in *Opposing* and *Own* conditions respectively) included any Weaken-own units. A GLMM using the Poisson distribution could not be executed, given means of zero in many cases.

Support-other usage. Support-other was the other form infrequently shown (less than 10%) in Year 1. In Year 2, however, 53% of participants in the *Opposing* condition included it once or more, compared to 9% in the *Own* condition participants did so. A significant increase was observed over time (from T1 to T3) in the percentage of *Opposing* condition students showing Support-other usage, as seen in Table 5 (Cochran's Q, $p < .001$), and from Year 1 to Year 2 (T2 vs T3), in particular, (McNemar, $p = .004$). A GLMM using the Poisson distribution showed condition differences, evident in the results of the overall model, $F(3, 102) = 6.048$, $p = .001$. The interaction between time and condition was also significant, $F(1, 102) = 12.995$, $p < .001$. The fixed effect of time as well as the fixed effect of condition were significant; fixed effect of time, $F(2, 102) = 9.311$, $p < .001$; fixed effect of condition $F(1, 102) = 14.309$, $p < .001$. Pairwise contrasts showed an increase over time in the *Opposing* condition, comparing T3 ($M = 1.053$, $SD = 1.268$, 95% CI = 0.441, 1.664) to both T2 ($M = 0.048$, $SD = 0.218$, 95% CI = -0.052, 0.147) and T1 ($M = 0$). No significant change was observed in the *Own* condition.

In the non-discourse topic, 26% of *Opposing* condition participants and 5% of *Own* condition participants included Supporting-other units in their essays. *Opposing* condition students showed a significant increase over time ($p = .007$, Cochran's Q), whereas no significant increase occurred in the *Own* condition (see Table 6). A GLMM using the Poisson distribution could not be executed, given a mean of zero in many cases.

Table 5

Table 6

Discussion

Caution is indicated in the comparison of Year 1 and Year 2 findings, as the Year 2 sample is reduced in size and the discussion topic differed by year. The results nevertheless are encouraging. The longitudinal findings for the reduced sample reported in Study 2 show that the gains observed in Year 1, after engagement in the intervention for one year, continued to accrue with a second year of engagement in participants' respective conditions, either equaling or surpassing Year 1 levels.

Furthermore, the differences between conditions became more pronounced in Year 2, most notably in the non-discourse topic. The percentage of *Opposing* condition participants including Weaken-other units at the end of Year 2 was twice that of the corresponding percentage observed at the end of Year 1, while no improvement over time was observed in the percentage of *Own* condition participants who included Weaken-other units, remaining at the low level of 5%. Furthermore, at Year 2 condition differences were also observed in frequencies of Weaken-other units, with *Opposing* condition students outperforming *Own* condition students, a finding observed only at Year 2. A similar condition difference appeared at Year 2 for Support-other statements. Over half of Study 2 participants in the *Opposing* condition included Support-other units in their essays on the discourse topic and a quarter of them did so in the non-discourse topic, compared to less than 10% of *Own* condition participants who included Support-other units for either the discourse or non-discourse topic.

In sum, both Study 1 hypotheses, of condition differences and skill transfer, as well as confirmation of the effectiveness of the intervention, received further support from Study 2.

General Discussion

We addressed our two research questions, one pertaining to condition effects and the other to transfer effects, by examining their manifestations in participants' post-discourse essays. We begin, however, with the Study 1 choice data from the discourse sessions themselves, as these are revealing in suggesting how the *Opposing* and *Own* condition participants' behavior during the discourse may have led to differences in their post-discourse essays.

Dialogs in the *Opposing* condition were conducted with opponents who held a contrasting view on the topic. Dialogs in the *Own* condition were conducted with those holding the same view. Despite equivalence of purpose, available information to use as evidence, and instructions across conditions, this difference influenced information-seeking behavior during the dialogs, as reflected in the qualitative data that were examined. Those whose dialogs began and continued only with those holding an opposing position sought information related to the opponents' position, which they were aware they would have to address and counter. Those in the *Own* condition, whose dialogs were confined to the like-minded, were also aware that they were going to engage in debate with opponents. However, the fact that this engagement was only anticipated apparently gave it less force and participants in the *Own* condition preferred at least initially to gain information that would help them build their own case (Table 4).

We turn now to the main data from both studies, the post-dialog essays. Wherever condition effects appeared they favored the *Opposing* condition. Consistent effects of time establish that a majority of participants in both conditions were able to understand and make use of the new information they accessed in constructing arguments that appeared in their final essays. Only those in the *Opposing* condition, however, achieved a near 100% success rate in constructing at least one idea unit that brings to bear relevant evidence pertaining to the claim

CONTEMPLATING THE OPPOSITION

made in that idea unit. These findings are consistent with other empirical research in which individuals engaged in argumentation with peers holding opposing views (Hemberger et al., 2017; Iordanou & Constantinou, 2015; Kuhn, Hemberger, & Khait, 2016; Villarroel, Felton, & Garcia-Mila, 2016).

While most functional evidence-based units across time and condition unsurprisingly functioned to support the writer's claim, units functioning to weaken the opponent's claims were not universal in either condition but became more prevalent across time in both conditions. Least prevalent overall, however, were evidence-based claims that were not concordant with the writer's position, serving either to weaken the writer's position or support the opposing position. Here also, *Opposing* condition participants were more likely to achieve mention of these, with a quarter of them (16) in Study 1 showing one of them, vs. 6% (4) in the *Own* condition. At Year 2, a majority of the *Opposing* condition participants achieved one or both, vs. 5% in the *Own* condition.

What accounts for the *Opposing* group's superiority? One could claim that their advantage lay in the fact that they had heard well-developed arguments expressed by their opponents, while the *Own* condition participants had to construct such arguments on their own. Making this less likely to be a sufficient explanation is the fact that the new information each side received was simple and straightforward in how it might be used to support a particular claim. Hence, participants in either condition should have experienced little difficulty in constructing such arguments among themselves, without hearing them expressed by someone from the opposing side, and they all had the same motive to do so as did those in the *Opposing* condition.

CONTEMPLATING THE OPPOSITION

Also needing to be considered in seeking an explanation of the condition effects is the fact that some participants showed transfer of their argument skills to a new topic, but such transfer was largely confined to those in the *Opposing* condition, who had been exposed to opposing peers' arguments and counterarguments on the main topic. This occurred even though information to use as evidence for the non-discourse topic was equally present across conditions and equally available at both times (in contrast to the discourse topic, where the information accrued only gradually during the discourse sessions). This transfer to a non-discourse topic was modest but increased with continued engagement and is therefore noteworthy, especially as the explanation for it is not obvious.

Why did this transfer of argument skill occur and why did the *Opposing* group show an advantage in this regard? The explanation we propose is that the *Opposing* group during their discourse were learning and practicing something about argumentation itself that some were then able to apply to a new topic – something as simple perhaps as “pay close attention to the arguments of the other side,” a hallmark of skilled argumentation. Accompanying this orientation is the recognition that the other is reasoning from a perspective that differs from one's own but may have as much coherence and therefore warrant attention and respect. Embodiment of this contrasting perspective in another person, as did our *Opposing* condition, may have supported this recognition. The different strategy patterns by condition that appeared in information choice also support this interpretation. It is also consistent with reports in the argumentation literature of gains in epistemological understanding of argumentation following encounters with diverse views (Barzilai & Ka'adan, 2017; Fisher, Knobe, Strickland, & Keil, 2017; Kienhues, Stadtler, & Bromme, 2011; Iordanou, 2016; Zavala & Kuhn, 2017; Kuhn et al., 2013).

CONTEMPLATING THE OPPOSITION

Our two research questions thus yield affirmative answers. Moreover, the effects they show are cumulative. Embodiment of a contrasting perspective in the form of an actual other person benefitted both learning about the topic, as reflected in the number and range of idea units that topic essays contained, and learning about argument, as reflected in gains manifested in essays on a new topic. These findings thus add support to the claim that the type of extended argumentation engagement and practice that we have employed, in which participants gain dense practice in direct peer-to-peer argumentation (Kuhn, 2018), can achieve both knowledge gains and skill gains during the course of the same activity (Iordanou et al., 2019). Such gains, moreover, are identifiable in the individual written argument that discourse serves as a bridge to, as has been demonstrated here. The effectiveness of the approach, we believe, is attributable to deep engagement with a topic and dense, targeted engagement with multiple, successive discourse partners (Kuhn, 2019).

The condition effect we designed the present study to investigate is not an original idea. It can be traced as far back as J.S. Mill (1859/1996), who expressed it quite precisely:

Nor is it enough that he should hear the arguments of adversaries from his own teachers, presented as they state them, and accompanied by what they offer as refutations. This is not the way to do justice to the arguments, or bring them into real contact with his own mind. He must be able to hear them from persons who actually believe them; who defend them in earnest, and do their very utmost for them.

Still, Mill was not addressing the transfer and generalizability of argumentation skills. We therefore must ask: How far can we generalize the present results from young adolescents’

CONTEMPLATING THE OPPOSITION

discourse on a non-threatening topic to adults discussing deeply-held, emotion-laden positions?

Only further investigation can say for sure. The present results do bear, however, on debates regarding how hard we should try to get people with different views to talk to one another (Hodson, 2011; Kuhn & Iordanou, in press). Passive listening to opposing views has shown scarce effect (Lao & Kuhn, 2002), whereas dense discourse with like-minded others may enrich thinking about a topic (Kuhn et al., 2019). Imagining positive contact with an out-group may enhance confidence in interacting with them (Crisp et al., 2011; Stathi et al., 2011). But Mill may have had it right that only getting to the other end of the continuum defined by genuine engagement will suffice. The fact that the present study has detected a difference in the context of a very affectively restricted exchange of views makes further investigation worthwhile.

References

- Andriessen, J., Baker, M., & Suthers, D. (Eds.). (2003). *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments*. Dordrecht: Kluwer.
- Asterhan, C., & Schwarz, B. (2016). Argumentation for learning: Well-trodden paths and unexplored territories. *Educational Psychologist*, 51(2), 164-187.
- Baron, J. (1994). *Thinking and deciding* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Barzilai, S., & Ka'adan I. (2017). Learning to integrate divergent information sources: The interplay of epistemic cognition and epistemic metacognition. *Metacognition and Learning*, 12(2), 193–232.
- Crisp, R. J., Birtel, M., & Meleady, R. (2011). Mental simulations of social thought and action: Trivial tasks or tools for transforming social policy. *Current directions in psychological science*, 20(4), 261–264.
- Crowell, A., & Kuhn, D. (2014). Developing dialogic argumentation skills: A 3-year intervention study. *Journal of Cognition and Development*, 15(2), 363-381.
- Fisher, M., Knobe, J., Strickland, B., & Keil, F. (2017). The influence of social interaction on intuitions of objectivity and subjectivity. *Cognitive Science*, 41(4), 1119-1134.
- Hemberger, L., Kuhn, D., Matos, F., & Shi, Y. (2017). A dialogic path to evidence-based argumentative writing. *Journal of the Learning Sciences*, 26(4), 575-607.
- Hodson, G. (2011). Do ideologically intolerant people benefit from intergroup contact? *Current Directions in Psychological Science*, 20(3), 154-159.
- Hodson, G., Crisp, R., Meleady, R., & Earle, M. (2018). Intergroup contact as an agent of cognitive liberalization. *Perspectives on Psychological Science*, 13(5), 523–548.

CONTEMPLATING THE OPPOSITION

- Iordanou, K. (2013). Developing face-to-face argumentation skills: Does arguing on the computer help? *Journal of Cognition and Development*. 14(2), 292-320.
- Iordanou, K. (2016). Developing epistemological understanding through argumentation in scientific and social domains. *Zeitschrift für Pädagogische Psychologie*. 30(2-3), 109-119.
- Iordanou, K. & Constantinou, C. P. (2015). Supporting use of evidence in argumentation through practice in argumentation and reflection in the context of SOCRATES learning environment. *Science Education*, 99(2), 282–311.
- Iordanou, K., Kendeou, P., & Beker, K. (2016). Argumentative reasoning. In W. Sandoval, J. Greene, & I. Bråten. (Eds.), *Handbook of epistemic cognition* (pp.39-53). New York: Routledge.
- Iordanou, K., Kuhn, D., Matos, F., Shi, Y., & Hemberger, L. (2019). Learning by arguing. *Learning and Instruction*. 63, 101207.
- Kienhues, D., Stadtler, M., & Bromme, R. (2011). Dealing with conflicting or consistent medical information on the Web: When expert information breeds laypersons' doubts about experts. *Learning and Instruction*, 21(2), 193–204.
- Kuhn, D. (2018). *Building our best future: Thinking critically about ourselves and our world*. Wessex Press.
- Kuhn, D. (2019). Critical thinking as discourse. *Human Development*, 62(3), 146-164.
- Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science*, 22(4), 545-552.
- Kuhn, D., Floyd, D., Yaksick, P., Halpern, M., & Ricks, W. (2019). How does discourse among like-minded individuals affect their thinking about a complex issue? *Thinking and Reasoning*. 25(3), 365-382. DOI: 10.1080/13546783.2018.1532460

CONTEMPLATING THE OPPOSITION

- Kuhn, D., Hemberger, L., & Khait, V. (2016). Tracing the development of argumentative writing in a discourse-rich context. *Written Communication*, 33(1), 92–121.
- Kuhn, D., & Iordanou, K. (in press). Why do people argue past one another rather than with one another? In N. Ballantyne & D. Dunning (Eds.), *Reason, bias, and inquiry: New perspectives from the crossroads of epistemology and psychology*. New York: Oxford University Press.
- Kuhn, D., & Lao, J. (1996). Effects of evidence on attitudes: Is polarization the norm? *Psychological Science*, 7(2), 115-120.
- Kuhn, D., Zillmer, N., Crowell, A., & Zavala, J. (2013). Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentative competence. *Cognition & Instruction*, 31(4), 456-496.
- Lao, J., & Kuhn, D. (2002). Cognitive engagement and attitude development. *Cognitive Development*, 17(2), 1203-1217.
- Lukianoff, G., & Haidt, J. (2018). *The coddling of the American mind*. New York: Penguin.
- Macagno, F. (2016). Argument relevance and structure: Assessing and developing students' uses of evidence. *International Journal of Educational Research*, 79(1), 180-194.
- Maier, J., Richter, T., Nauroth, P., & Gollwitzer, M. (2018). For me or for them: how in-group identification and beliefs influence the comprehension of controversial texts. *Journal of Research in Reading*. 41(1), 48–65.
- Mill, J.S. (1859/1996). On liberty. In D. Wootton (Ed.), *Modern political thought: Readings from Machiavelli to Nietzsche*. Indianapolis/Cambridge: Hackett.
- Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly*, 46(3), 273-304.

CONTEMPLATING THE OPPOSITION

- O'Connor, C. & Snow, C. (2018). Classroom discourse: What do we need to know for research and for practice? In M. Schober, D. Rapp, & M.A. Britt (Eds.), *Routledge Handbook of discourse processes* (pp. 315-342). New York: Routledge.
- Rapanta, C., Garcia-Mila, M., & Gilabert, S. (2013). What is meant by argumentative competence? An integrative review of methods of analysis and assessment in education. *Review of Educational Research*, 83(4), 483-520.
- Resnick, L., Asterhan, C. & Clarke, S. (Eds.) (2015). *Socializing intelligence through academic talk and dialogue*. Washington: American Educational Research Association.
- Resnick, L., Asterhan, C., Clarke, S., & Schantz, F. (2018). Next generation research in dialogic learning. In G. Hall, L. Quinn, & D. Gollnick (Eds.), *Wiley Handbook of teaching and learning* (pp. 323-338). New York: Wiley.
- Shi, Y., Matos, F., & Kuhn, D. (2019). Dialog as a Bridge to argumentative writing. *Journal of Writing Research* 11(1), 107-29.
- Stathi, S., Crisp, R. J., & Hogg, M. A. (2011). Imagined inter-group contact enables member-to-group generalization. *Group dynamics: Theory, research and practice*, 15(3), 275–284.
- Villarroel, C., Felton, M., & Garcia-Mila, M. (2016). Arguing against confirmation bias: The effect of argumentative discourse goals on the use of disconfirming evidence in written argument. *International Journal of Educational Research*, 79(1), 167-179.
- Walton, D. N. (1989). Dialogue theory for critical thinking. *Argumentation*, 3(2), 169-184.
- Zavala, J., & Kuhn, D. (2017). Solitary discourse is a productive activity. *Psychological Science*, 28(5), 578-586.

CONTEMPLATING THE OPPOSITION

Table 1

Types of Argument Functions

Type	Definition	Example*
Support my own	A statement serving to support one's own position (M+)	It's more economical for a household to produce electricity using solar energy.
Weaken other	A statement serving to critique and thereby weaken the opposing position (O-)	Many of the areas that have been explored for natural gas are natural areas and the development of these areas can have negative effects on the natural environment and the wildlife in the area.
Support other	A statement serving to acknowledge strengths of the opposing position (O+)	Natural gas offers electricity to industry day and night, but ...
Weaken my own	A statement serving to acknowledge weaknesses of one's own position (M-)	Someone could say it's not feasible to produce electricity at night using the solar energy, but ...

*The examples are from a participant preferring the solar energy position.

CONTEMPLATING THE OPPOSITION

Table 2. Number (and percentages) of Study 1 Participants who Used Functional Units, Weaken-other, Weaken-own and Support-other Evidence Units at Least Once in the Discourse-topic Essay

	<i>Opposing Condition</i>		<i>Own Condition</i>	
	Initial Assess.	Final Assess.	Initial Assess.	Final Assess.
	N = 65	N = 65	N = 65	N = 65
Functional Units	44 (68%)	61 (94%)**	41 (63%)	40 (62%)
Weaken-Other	12 (19%)	40 (62%)**	12 (19%)	41 (63%)**
Weaken-Own	1 (2%)	10 (15%)*	1 (2%)	2 (3%)
Support-Other	1 (2%)	6 (9%)	1 (2%)	2 (3%)

*Significant change, $p = .004$, McNemar Test; **Significant change, $p < .001$, McNemar Test.

CONTEMPLATING THE OPPOSITION

Table 3. Number (and percentages) of Study 1 Participants who Used Functional Units, Weaken-Other, Weaken-Own and Support-Other Evidence Units at Least Once in the Non-Discourse-topic Essay

	<i>Opposing Condition</i>		<i>Own Condition</i>	
	Initial Assess. N = 63	Final Assess. N = 60	Initial Assess. N = 64	Final Assess. N = 64
Functional Units	51 (81%)	58 (97%)*	45 (70%)	55 (86%)**
Weaken Other	4 (6%)	11 (18%)*	3 (5%)	3 (5%)
Weaken Own	0	1 (2%)	0	0
Support Other	0	1 (2%)	2 (3%)	2 (3%)

*Significant change, $p = .016$, McNemar Test, ** Significant change, $p = .002$, McNemar Test.

CONTEMPLATING THE OPPOSITION

Table 4

Frequencies of Information Choice Type Across Sessions by Condition

	<i>Opposing</i> Condition			<i>Own</i> Condition		
	Predominantly Own	Mixed Own & Other	Predominantly Other	Predominantly Own	Mixed Own & Other	Predominantly Other
Initial (first 4) sessions	1	1	4	3	3	0
Later (last 4) sessions	5	0	1	1	2	3

CONTEMPLATING THE OPPOSITION

Table 5. Number (and Percentages) of Study 2 Sample who had at Least One Weaken-other, Weaken-own and Support-other Units in Essays on the Discourse Topic

	<i>Opposing Condition</i>			<i>Own Condition</i>		
	Time			Time		
	T1 (N = 21)	T2 (N = 21)	T3 (N = 19)	T1 (N = 23)	T2 (N = 21)	T3 (N = 23)
Weaken- other	3 (14%)	17 (81%)	15 (79%) **	3 (13%)	12 (57%)	12 (52%) **
Weaken- own	0	4 (19%)	10 (53%)*	0	1 (5%)	2 (9%)
Support- Other	0	1 (5%)	10 (53%) **	0	0	2 (9%)

*Significant change over time, $p = .001$, Cochran's Q

** Significant change over time, $p < .001$, Cochran's Q

CONTEMPLATING THE OPPOSITION

Table 6. Number (and Percentages) of Study 2 Sample who had at Least One Weaken-other, Weaken-own and Support-other Units in their Essays on the Non-discourse Topic

	<i>Opposing</i> Condition			<i>Own</i> Condition		
	Time			Time		
	T1 (N=20)	T2 (N = 20)	T3 (N = 19)	T1 (N = 22)	T2 (N = 22)	T3 (N = 21)
Weaken- other	0	2 (10%)	6 (32%)*	0	1 (5%)	1 (5%)
Weaken- own	0	0	2 (11%)	0	1 (5%)	1 (5%)
Support- Other	0	0	5 (26%)**	0	0	1 (5%)

*Significant change over time, $p = .009$, Cochran's Q

** Significant change over time, $p = .007$, Cochran's Q

CONTEMPLATING THE OPPOSITION

Figure 1. Mean usage of functional units for the two conditions in the non-discourse topic.

