The role of language in students’ justifications of chemical phenomena

Jacky M. Deng\textsuperscript{a}, Malek Rahmani\textsuperscript{a}, and Alison B. Flynn\textsuperscript{a*}

\textsuperscript{a}Department of Chemistry and Biomolecular Sciences, Faculty of Science, University of Ottawa, Ottawa, Ontario, Canada

Abstract
Making decisions and constructing arguments with scientific evidence and reasoning are essential skills for all members of society, especially in a world facing complex socioscientific issues (climate change, global pandemics, etc.). Argumentation is a complex linguistic practice, but little is known about how students from diverse language backgrounds engage in argumentation. The goal of this study was to identify how students’ English language proficiency/history was associated with the reasoning demonstrated in their written arguments. We found that students with lower English proficiency and less English history produced fewer causal responses compared to students with higher English language proficiency and history. Follow-up interviews with fifteen participants revealed that students’ comfort communicating in English on assessments depended on a combination of general and academic language experiences. Findings suggest a need to identify what barriers students from diverse language backgrounds encounter during argumentation to ensure students from all language backgrounds have equitable opportunities to demonstrate their abilities.

Introduction
In a world facing complex socioscientific issues (e.g., climate change, viral pandemics, misinformation), citizens need to be able to build arguments and reason from scientific evidence (National Research Council, 2013; Organisation for Economic
Cooperation and Development, 2006; Social Sciences and Humanities Research Council, 2018; United Nations, 2015). Students can develop scientific argumentation skills through classroom activities and assessments (e.g., asking students to justify why a phenomenon occurs), but the linguistic skills required to express clear arguments can be a challenge for students (McNeill & Berland, 2010; McNeill & Krajcik, 2008b, 2008a; McNeill, Lizotte, Krajcik, & Marx, 2006). This can be especially challenging for students who learned English-as-an-additional language (Eng+) who are expected to learn, use, and communicate abstract scientific concepts in academic language while they are simultaneously learning English (Abedi, Zhang, Rowe, & Lee, 2020; Afitska & Heaton, 2019; C. Buxton et al., 2014a; Curtis & Millar, 1988). Language is foundational to teaching and assessment; it plays in participating in science and constructing meaning. Students’ challenges with language may also be exacerbated in disciplines like chemistry in which terms like “stability” and “resonate” have different meanings from their everyday definitions (Childs, Markic, & Ryan, 2015; Uhl Chamot, 1995).

Eng+ students face unique linguistic challenges with knowing, doing, and talking science (O. Lee & Fradd, 1998). Knowing science means making meaning of scientific knowledge by connecting new information to previous knowledge. This can be a challenge for Eng+ students, as they may have learned prerequisite concepts in a language or culture different from English, making it more challenging to identify and establish coherent connections between new information and their previous knowledge. Doing science involves scientific practices, such as engaging in inquiry, proposing arguments and explanations, and interpreting and identifying evidence to make sense of the world (O. Lee & Fradd, 1998; O. Lee, Quinn, & Valdés, 2013;
National Research Council, 2013). Many of these practices require complex language functions and academic traditions that require sufficient levels of language proficiency within a given discipline. *Talking* science means communicating in the language of science: academic English (Lemke, 1990). Eng+ students may struggle to learn or demonstrate the expected communication patterns of academic English. For example, in contexts like exams where students may be asked to articulate an explanation or argument for why a phenomenon occurs, Eng+ students may engage in additional cognitive operations compared to English-as-a-first (Eng1st) language peers students in order to interpret English questions and then generate a written output in English (Abedi, 2002, 2015; Abedi, Hofstetter, & Lord, 2004; Abedi & Lord, 2001; C. Buxton et al., 2014b; González-Howard & McNeill, 2016; E. N. Lee, 2018; E. N. Lee & Orgill, 2021; Solano-Flores & Trumbull, 2003) (Figure 1).

Figure 1: An outline of processes students use to solve problems, with additional processes that Eng+ students may engage in highlighted in red. Adapted from E.N. Lee (2018).

The abovementioned linguistic challenges faced by Eng+ students can manifest into inequities in how Eng+ students are assessed and evaluated in science classrooms and educational research. Eng+ students’ achievement in STEM has been found to be associated with language ability, regardless of students’ abilities to grasp and apply scientific concepts (Curtis & Millar, 1988; Maerten-Rivera, Myers, Lee, & Penfield, 2010; Solano-Flores & Trumbull, 2003). Eng+ students’ written responses can score
lower on evaluations than those of English-as-a-first language (Eng1st) students due to evaluators’ implicit bias (Huang, 2008; Lindsey & Crusan, 2011; Milnes & Cheng, 2008) and students’ challenges communicating their knowledge (Lyon, Bunch, & Shaw, 2012; Swanson, Bianchini, & Lee, 2014; Wolf et al., 2008). As Kieffer et al. (2009) write: “language plays an integral role in most, if not all, academic learning, any test of academic achievement is also, to some degree, a test of language ability” (Kieffer, Lesaux, Rivera, & Francis, 2009).

Neglecting the issues facing Eng+ students—who are already largely overlooked in postsecondary science courses (Kanno & Cromley, 2013)—can have problematic downstream consequences, including increased attrition of students from a diversity of linguistic, cultural, and ethnic backgrounds, and curbed development of key skills, such as scientific argumentation (Abedi, 2002; Afitska & Heaton, 2019; Allensworth, 2005; C. A. Buxton, Salinas, Mahotiere, Lee, & Secada, 2015; C. Buxton et al., 2014a; Chin, 2010; Maerten-Rivera et al., 2010; McNeil, 2005; National Academies of Science Engineering and Medicine, 2018; Pyburn, Pazicni, Benassi, & Tappin, 2013; Solano-Flores & Trumbull, 2003; Sotelo-Dynega, Ortiz, Flanagan, & Chaplin, 2013; Wolf et al., 2008; Yore et al., 2004). Eng+ students form a significant and increasing population of students in science classrooms around the world. There are over 1,250,000 international students studying in North America, the majority of whom come from nations where English is not a primary language (Project Atlas, 2019). In the United States (US), 4.6 million Eng+ students make up about 9.4% of all classrooms, with 1.1 million Eng+ students in US postsecondary institutions (National Center for Education Statistics, 2017). Of STEM graduate students in the US, 38% are temporary
visa holders from non-English nations, a number that has steadily increased since 2000 (National Science Foundation, 2020). In the United Kingdom, the Eng+ student population has increased 3.7% since 2012 (21.2% of all students) (Oxley & de Cat, 2019).

In this study, we investigated how English language proficiency and history was associated with the reasoning in students’ written chemistry arguments. Much of the current work on language and learning has been conducted in K–12 mathematics, biology, or general science contexts, with few investigations in chemistry (Pyburn et al., 2013) and post-secondary contexts (LaCosse, Canning, Bowman, Murphy, & Logel, 2020), despite chemistry’s dependence on Anglophonic nomenclature, models, and symbols (Childs et al., 2015). Educational researchers have studied students’ scientific reasoning and argumentation skills in various contexts and disciplines (Becker, Noyes, & Cooper, 2016; Bodé, Deng, & Flynn, 2019; Cooper, Kouyoumdjian, & Underwood, 2016; Deng & Flynn, 2021; Moon, Moeller, Gere, & Shultz, 2019; Moon, Stanford, Cole, & Towns, 2016; Osborne & Patterson, 2011; Russ, Scherr, Hammer, & Mikeska, 2008; Watts et al., 2020; Weinrich & Talanquer, 2016); however, much of this work has relied on analyses of arguments constructed in a single language (usually English), without considering the role of language in these types of investigations.

Goals and research questions

Our research questions were:

(1) What differences in reasoning might there be in arguments constructed by postsecondary science students who have learned English-as-an-additional language (Eng+) and English-as-a-first language (Eng1st)?
(2) In what ways might English language proficiency and history be correlated with the reasoning postsecondary science students employ when constructing an argument in chemistry?

Methods and Materials

Context, recruitment, and sample

The study was conducted at the University of Ottawa, a bilingual institution in which students have the option to complete coursework in either English or French. All students from non-English backgrounds must complete an English-language test (e.g., TOEFL) to enrol in courses taught in English at the University. With a diverse student population, many students speak a language other than English (University of Ottawa, 2021). To our knowledge, this is the first study of post-secondary students who may be completing their studies in either English, French, or both.

We asked professors teaching chemistry courses at the University to forward a recruitment text to students in their courses via email. The final sample population was composed of 166 participants who were students in chemistry courses across all four academic years. The institution’s Office of Research Ethics and Integrity approved the project prior to recruitment (H11-18-1363).

Instruments and data collection procedure

In an online format, participants were asked to construct an argument in response to a social media post related to ocean acidification using chemical equilibria data (Figure 2), which was based on a prompt used in a previous study on students’ argumentation in chemistry (Moon et al., 2019). Participants were also asked to report demographic information, including whether they identified as Eng+ or Eng1st, the
primary language in which they had learned chemistry, and their English language proficiency and history using the Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian, Blumenfeld, & Kaushanskaya, 2007). The instruments and evidence for validity and reliability for measures can be found in the Supplementary Materials (SM).

Figure 2: Participants were asked to construct an argument to the social media post (left) using relevant chemical equilibria (right).

We conducted standardized, open-ended interviews with a subset of participants \( n = 15 \) to gain additional insight into participants’ perceptions of how language contributes to their experiences when communicating in science and chemistry education.

**Data analysis**

We analysed participants’ arguments to identify the claim, evidence used, and mode of reasoning: descriptive, relational, linear causal, or multicomponent causal.
This analytical framework has been used to successfully capture progressions in chemistry reasoning and that it requires demonstrating increasingly sophisticated linguistic complexity as they progress from descriptive to causal modes of reasoning. *Descriptive arguments* list or give features and/or the properties of entities (e.g., reactants, products) without establishing connections. *Relational arguments* include connections between properties of the entities and their activities, but these relationships are discussed in a correlative fashion (i.e., lack causality). *Causal arguments*—linear and multi-component—include all features of a relational argument and additionally contain cause-and-effect relationships between the relevant properties of the entities and their activities.

An initial coding protocol was developed by author JD using the entire dataset (N = 166). Inter-rater reliability (IRR) was measured using Krippendorf’s α to mitigate inflation of agreement due to chance agreements between raters (Krippendorff, 1970). The IRR process began with co-authors JD and MR independently coding 15% (n = 26) of the full data set with the initial codebook. Krippendorff’s α was calculated to be 0.372. After discussing disagreements, authors JD and MR modified the codebook and conducted a second round of coding by coding another 15% of the full data set with the modified codebook. After this second round, Krippendorff’s α was calculated to be 0.908. Throughout the process, participants’ arguments were not organized in terms of Eng+ and Eng1st groups. This was done to limit bias, ensuring that coders were blind to
which arguments were constructed by Eng+ and Eng1st students. No arguments demonstrated multi-component causal reasoning in the dataset and so that mode of reasoning does not appear in the codebook or subsequent description of findings.

Table 1: Modes of reasoning used to characterize the reasoning demonstrated in participants’ arguments.

<table>
<thead>
<tr>
<th>Mode of reasoning</th>
<th>Definitions, adapted to the prompt</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive</td>
<td>Argument describes explicit features of the graph or equations (e.g., as atmospheric CO₂ concentration goes up, pH goes down).</td>
<td>Based on the equations, H₂O⁺ is generated. Based on the graph, CO₂ goes up as pH goes down.</td>
</tr>
<tr>
<td>Relational</td>
<td>Argument identifies that increased CO₂ concentrations (atmospheric or aqueous) are correlated with H⁺ being generated, using the equations. Identifies that increased H⁺ is correlated with decreasing seawater pH.</td>
<td>Based on the graph, CO₂ goes up as pH goes down. As is shown in the equations, increasing CO₂ produces more H₂O⁺ ions. Based on the equations, H₂O⁺ is generated. Increasing H₂O⁺ makes the seawater more acidic, so pH goes down.</td>
</tr>
<tr>
<td>Linear Causal</td>
<td>Argument uses Le Chatelier’s principal or pKₜ values to justify why atmospheric CO₂ concentrations influence the directions of the chemical equilibria and connects back to the impact that has on seawater pH.</td>
<td>Based on the graph, CO₂ goes up as pH goes down. Due to Le Chatelier’s principle, increasing CO₂ produces carbonic acid, which leads to increased concentration of H₂O⁺ ions. This results in increased acidification of the seawater, and thus a lower pH.</td>
</tr>
</tbody>
</table>

English language profiles were created for each participant by using respondents’ LEAP-Q data developing to develop English language proficiency and history indices (Figure S1) (Krizman, Skoe, Marian, & Kraus, 2014; Marian, Chabal, Bartolotti, Bradley, & Hernandez, 2014; Reichle & Birdsong, 2014). Statistical tests were used to evaluate how participants’ language backgrounds were associated with their demonstrated mode of reasoning. All statistical tests were conducted in SPSS (details in the SM). Fischer’s exact test was used to test independence of the modes of reasoning between Eng+ and Eng1st responses, followed by post-hoc pairwise Z-tests with Bonferroni adjusted p-values for multiple comparisons. Cluster analysis was conducted based on
two clustering variables: relative English language history and relative English language proficiency. After evaluating parameters for collinearity (Table S2), full factorial multinomial logistic regression was used to evaluate the association between English language proficiency and history (independent covariates) and mode of reasoning (dependent categorical output). Academic year was included as a control variable (categorical predictor). Findings from the first regression analysis prompted a second logistic regression, with mode of reasoning as the dependent variable and (a) time in English language country, (b) time in English language family, and (c) time in English language school/workplace as independent predictors. Lastly, we investigated the relationships between language group (Eng+, Eng1st), language of chemistry instruction (English, French), and reasoning demonstrated in participant’s arguments (descriptive, relational, causal).

Interviews were transcribed and analysed thematically (Cresswell, 2012; Lincoln & Guba, 1985), with attention to: (a) how Eng+ participants perceived the role of language in their own learning experiences and (b) how Eng+ and Eng1st participants perceived the role of language in others’ learning experiences (peers, friends, classmates).

**Results**

**Eng+ participants constructed fewer causal arguments than Eng+ participants**

Eng+ participants constructed fewer causal arguments than Eng1st participants and instead gave more descriptive and relational arguments, \( \chi^2 (1, N = 166) = 11.09, p < 0.01, \phi = 0.253 \) (Figure 3); in this analysis, Eng1st and Eng+ participants were treated as dichotomous groups. Eng+ arguments commonly described what was occurring
(descriptive) and/or how the data presented to them were connected (relational). For example Participant 28 acknowledged how the creation of \( \text{H}_3\text{O}^+ \) could be related to decreasing pH, but did not further elaborate on how or why these ideas are related:

I disagree with Sam. The equations show clearly a relationship between the presence of \( \text{CO}_2 \) and the formation of \( \text{H}_3\text{O}^- \). \( \text{H}_3\text{O}^+ \) makes the pH of the water to go down.

In contrast, Eng1st arguments more frequently described why the data presented were connected and relevant to each other, often citing Le Chatelier’s principle to support their argument (linear causal). For example, Participant 38:

I disagree with Sam. According to Le Chatelier’s principle as \( \text{CO}_2 \) in the atmosphere increases, \( \text{CO}_2 \) in the ocean must also increase. This means \( \text{CO}_2 \) in the first line of the series of equilibria will be pushed towards the carbonic acid product. This in turn will interact with the water producing acidic hydronium ions/hydrogen ions. The conjugate base produced from the second line of the chemical equilibria can still lose a hydrogen to a surrounding water molecule, creating more acid species. The \( pK_a \) values indicate the acid strength (how much they dissociate in water), and both of the carbonic acid species are more acidic than water, contributing to the acidification of the ocean.
Participants with lower English history and proficiency scores produced the fewest causal arguments

We encountered a challenge with the Eng+ group in that participants were considered Eng+ both if they are relatively new to English and experienced with English-as-an-additional language. Heterogeneity in the Eng+ and Eng1st samples (identified in the LEAP-Q data) is a common challenge when categorizing individuals into language groups (O. Lee & Stephens, 2020; National Academies of Science Engineering and Medicine, 2018). To better capture the heterogeneity between groups, we used cluster analysis using the LEAP-Q data. This analysis uncovered four language clusters (Figure 4).

Figure 4: Participants’ English proficiency and history, organized into four clusters using cluster analysis (N = 166).

Cluster 1 responses more frequently exhibited causal reasoning compared to Cluster 4, 

\[ \chi^2 (6, N = 166) = 14.02, p < 0.05, \phi = 0.284 \]
a). The proportion of causal reasoning decreased from Cluster 1 to Cluster 4, while the proportion of descriptive and relational reasoning increased.

Cluster 1 had the highest proportion of participants who identified as Eng1st (99%), $\chi^2(9, N = 166) = 94.97, p < 0.0001, \phi = 0.756$ (b). Within Cluster 1, 91% had learned chemistry in primarily English (EngCHM) and 8% having learned chemistry primarily in French (FrCHM). In contrast, Cluster 4 had the highest proportion of Eng+ participants (79%), with 42% having learned chemistry in English and 37% in French.

Higher English history was predictive of generating a causal argument

English language history was a significant predictor of exhibiting causal reasoning relative to the reference category (relational reasoning), controlling for academic year, Exp($B$) = 1.057, $p < 0.05$ (Table S3). That is, participants who had more experiences and time with English were more likely to produce a causal response compared to a relational response. More time in an English language family was a
significant predictor of exhibiting causal reasoning relative to relational reasoning, \( \text{Exp(B)} = 1.085, p < 0.05 \) (Table S4).

Eng+ participants who had learned chemistry primarily in French did not demonstrate causal reasoning at all in their arguments (Figure 6). The proportion of causal arguments increased as the alignment between language group and language of instruction increased: Eng+ participants in English chemistry (16.7% causal), Eng1st participants in French chemistry (37.0% causal), and Eng1st participants in English chemistry (44.1% causal).

![Figure 6: Intersection between language group (Eng+, Eng1st), primary language of chemistry instruction (EngCHM, FrCHM) and reasoning.](image)

**Preferred language for communicating in chemistry was connected to general and chemistry-specific language experiences**

Most participants preferred communicating chemistry in the language of chemistry instruction, regardless of their first language (Figure 7, with examples in Table S8).

Both Eng1st and Eng+ participants who had learned most chemistry in English \( (n = 5 \)
and \( n = 4 \), respectively) said they preferred English in chemistry contexts. They said it would be difficult for them to communicate the same ideas in their first and/or home languages. For example, Steph, a student who spoke Urdu and Punjabi as their first language at home, when asked if how they would communicate chemical ideas in these languages:

I would definitely use a lot of English words, just because I wouldn't know the words for the things I want to talk about in Urdu or Punjabi. So, you would definitely hear a lot of English...Like, the Urdu and Punjabi I know is more just conversational. Not really scientific.

Figure 7: Based on interview data, participants’ comfort with responding in English vs. French was connected both their general language background (Eng1st or Eng+) and their primary language of academic chemistry instruction (English or French).

Two participants who had learned chemistry in English but spoke French in social and home contexts said they could also communicate chemistry effectively in French given the bilingual context \( (n = 2) \). Those participants also had high English history and proficiency scores. They said they thought about chemistry in English and could
translate this knowledge into French if needed due to prior experiences learning some chemistry in French. For example, Jen stated:

I did my schooling all in French, all the way up to second year. And then like, I just went full English. My dad and his whole side of the family is English, so I'm pretty equally comfortable with both languages... I find that in courses like biology and stuff like that, my brain still functions in French because I took those courses in French. When I'm doing chemistry, my brain thinks in English, so it actually comes easier to me in English, because I have more experience...in French I think it might even be a bit harder, because I'd have to translate some of the words, but I mean, I wouldn't say that it'd be hard. I could still do it easily. But I would have to mentally translate.

Eng1st participants learning chemistry in French said they often thought about chemistry in French and would be most easily able to generate a response in French as in English ($n = 3$). They believed French allowed them to communicate more effectively and provide more detail than English. Melanie stated:

I've done my whole schooling in French, so all the terms and stuff are in French in my head... In French, instead of saying, "I would disagree with the Twitter user, it is not misleading", I would explain why it was not misleading, instead of just saying it. Because in English, I feel like I affirm more things instead of explaining why it's not as misleading. So, I would have probably went into the why a little bit more.

These findings were reinforced by interview participants’ LEAP-Q data. Participants who exhibited relatively higher levels of English history and proficiency stated that they were confident in their abilities to communicate chemistry in English ($n = 5$). In contrast, describing a need to mentally translate chemistry knowledge from French to English or vice versa was associated with lower English history ($n = 5$), while a preference French over English was associated with relatively lower scores of both
English history and proficiency ($n = 3$). The statistical significance of observations is limited by the small sample size and more research is needed to investigate the interplay of the various factors involved.

**Discussion**

*Supporting students’ reasoning abilities requires supporting both ability and expression*

Three main findings emerge from this work: (1) Eng+ participants produced fewer arguments with causal reasoning compared to Eng1st participants; (2) participants with lower English language proficiency and history produced fewer causal responses compared to participants with higher English language proficiency and history; and (3) participants often preferred communicating chemistry in the same language as they had learned it.

Taken together, the findings suggest that language ability is correlated with the reasoning demonstrated in students’ written assessments in chemistry, reinforcing previous findings about Eng+ students’ challenges generating written arguments and explanations of phenomena (Abedi, 2002, 2015; Abedi et al., 2004; Abedi & Lord, 2001; C. Buxton et al., 2014b; González-Howard & McNeill, 2016; E. N. Lee, 2018; E. N. Lee & Orgill, 2021; Solano-Flores & Trumbull, 2003).

However, our findings reveal that *students who are less proficient in English may still prefer to use English* for science communication. Students who are less proficient in English and/or consider themselves Eng+ and speak a non-English language at home may still prefer to communicate academic ideas in English based on their experiences...
using English in academic settings. This finding is reinforced by previous literature distinguishing between English for academic purposes and English for general/social purposes (Cummins, 1981; Flowerdew, 2019; Hyland, 2016; O. Lee & Fradd, 1998; Myles & Cheng, 2003). Certain linguistic academic accommodations, such as bilingual dictionaries, may not be helpful for Eng+ learners. Instead, other academic supports may be more useful to achieve equity; more research is needed.

Participants who had the fewest English experiences (both at home and in academic settings) produced the lowest proportion of causal arguments and felt less comfortable using English in an academic science setting. This finding suggests that students from majority non-English backgrounds may be the most disadvantaged when expected to communicate academic knowledge in English. In many contexts, science may be taught entirely in a non-English language; for example, in Québec, Canada, students have the option to complete the entirety of their science education in French (Les CÉGEPs du Quebec, 2019). Individuals from these backgrounds may have little experience in communicating in English, either in general or scientific contexts, consequently limiting their participation in science or making participation much more challenging (Cheng et al., 2019; Elnathan, 2021; Ortega, 2021). Scientists from Eng+ backgrounds can face unique financial and career costs in having to navigate a scientific enterprise dominated by English (Cheng et al., 2019; Elnathan, 2021; Hanauer & Englander, 2011; Hanauer, Sheridan, & Englander, 2019; Ortega, 2021; Ramírez-Castañeda, 2020).
Limitations

This study used convenience sampling so self-selection bias may have been present. Eng+ participants in this study may not be representative of the full diversity within the Eng+ student population, though we included statements in our recruitment texts to explicitly invite students from linguistically diverse backgrounds. Based on interview and LEAP-Q data, most of the Eng+ participants appeared to be long-term English learners and/or English/French bilingual students whose experiences may be distinct compared to English newcomers. Future endeavours involving Eng+ students could employ alternative recruitment methods that invite participation from a diverse cohort of the Eng+ population (Liamputtong, 2008).

The low-stakes context (i.e., voluntary questionnaire) may have contributed to participants responding differently than in higher-stakes contexts (e.g., an exam). For example, Eng+ students may have misinterpreted the expectations of the prompt/context, which may have shaped how they responded in the study. During the interviews, both Eng+ and Eng1st participants who demonstrated descriptive or relational reasoning stated that they would have either provided more detail or tailor their response to course expectations on an exam (Table S7). One Eng+ participant who produced a causal response explained how their self-perceived limitations with English often pushed them to provide greater detail and clarity on assessments to ensure they were being understood correctly. Despite this finding, we hesitate to make general claims about how Eng+ and Eng1st participants may have perceived the stakes differently given the limited sample size and lack of concrete evidence on how participants would respond in a higher-stakes context.
Other demographic factors may have contributed to how participants responded, including grade-point average, socioeconomic status, racial/ethnic background, immigrant/newcomer status, based on previous research (Afitska & Heaton, 2019; González-Howard & Suárez, 2021; LaCosse et al., 2020; O. Lee, 2005). The small sample size for the follow-up interviews limits the strength of the findings and the generalizability to other contexts. Nevertheless, important aspects of language were uncovered that merit further study.

**Conclusions and implications**

Constructing scientific arguments is an essential skill for members of society, especially in a world facing complex socioscientific issues. However, scientific argumentation is a complex linguistic practice and there is limited research on how language proficiency is related to individuals’ argumentation skills. Therefore, we investigated how undergraduate science students’ English language proficiency and history were associated with their reasoning in scientific arguments using prompt on ocean acidification.

We found that participants’ English language experiences, including their proficiency and history, were associated with the type of reasoning demonstrated. Fewer scientific arguments from English-as-an-additional language (Eng+) participants demonstrated causal reasoning compared to arguments from English-first language (Eng1st) participants. Participants with high English proficiency and history (Cluster 1) produced more arguments with causal reasoning than participants with low English proficiency and history (Cluster 4). Regression analyses revealed that English language history was a significant predictor of the reasoning exhibited in participants’ arguments.
In interviews, participants described how their language preferences depended on the contexts in which they had learned a particular language. Participants described a preference for using their primary language of chemistry instruction when engaging with chemistry content, even if the language of instruction was not their preferred language in general contexts. Participants who had the fewest English experiences (both out of and in school) felt the least comfortable using English in an academic science setting.

Participating in science requires being able to communicate disciplinary content knowledge in academic English, and language barriers faced by Eng+ may hinder their abilities to demonstrate both disciplinary content knowledge and scientific practices, such as argumentation. Implications may also extend to the state of diversity in science more generally—in later career settings, Eng+ professionals who have learned chemistry in non-English contexts may not be able to fully express themselves in English. That gap between ability and expression can not only hinder a person’s career (Cheng et al., 2019; Elnathan, 2021), but also limits that benefit that the scientific (or other) community could gain from their expertise (Amano, González-Varo, & Sutherland, 2016; Ramírez-Castañeda, 2020).

Future research may investigate these research questions in higher stakes settings (e.g., exams) and identify other variables that may be at play (e.g., sophistication of arguments in French settings). Studies could investigate the specific challenges that exist, the nature and timing of barriers, and the approaches Eng+ students use to succeed in English environments, and the effects of equity-intended interventions. The results of these studies could be used to better support Eng+ students through
interventions (e.g., educational opportunities) or training for educators (e.g., designing inclusive assessments, mitigating implicit bias when evaluating work).

Academic accommodations to support Eng+ students have been proposed but need further study, including providing students with more time to complete assessments, incorporating scaffolds and visual aids on assessment items, and low-stakes opportunities to develop language and argumentation skills (Abedi et al., 2004, 2020; Cikmaz, Fulmer, Yaman, & Hand, 2021; Francis, Rivera, Leseaux, Kieffer, & Rivera, 2006; E. N. Lee & Orgill, 2021; Ryoo & Bedell, 2019; Siegel, 2007). Others have suggested avenues focused on universal design (Thompson & Thurlow, 2002).

Education can maximize the experiences that Eng+ students already bring to the classroom, framing language backgrounds as assets and not deficits, as Eng+ students are a diverse group of learners who bring unique multilingual perspectives to the classroom. Moreover, persistent deficit-framing can limit Eng+ students’ engagement in science learning experiences (González-Howard & Suárez, 2021; O. Lee, 2021).

Creating a more equitable and inclusive society requires mitigating barriers and treating diversity as an asset to support participation of individuals from all cultures and experiences. By better understanding language-related barriers to learning in science, we can better support students and scientists throughout their careers, enhancing both equity and innovation in science (Bell, Villado, Lukasik, Belau, & Briggs, 2011; Harris, Mack, Bryant, Theobald, & Freeman, 2020; Hofstra et al., 2020; Impellizzeri & Coe, 2021; O. Lee et al., 2013; McKinsey & Company, 2015; Valantine, Collins, & Verma, 2015).
Acknowledgements

This work was supported by Canada’s Social Science and Humanities Research Council (SSHRC) under the Vanier Graduate Scholarship program.

References


Les CÉGEPs du Quebec. (2019). What is a CÉGEP?


Lyon, E. G., Bunch, G. C., & Shaw, J. M. (2012). Navigating the language demands of an


Wolf, M. K., Kao, J. C., Griffin, N., Herman, J., Bachman, P. L., Chang, S. M., &
