The role of language in students' justifications of scientific phenomena

Authors: Jacky M. Deng[1], Malek Rahmani[1], Alison B. Flynn[1]*

Affiliations:

[1] Department of Chemistry and Biomolecular Sciences, Faculty of Science, University of Ottawa, Ottawa, Ontario, Canada

*Corresponding author. Email: alison.flynn@uottawa.ca

Abstract: Constructing scientific arguments is an essential skill for members of society, especially in a world facing complex socioscientific issues. Educators can help students develop scientific argumentation skills; however, argumentation is a complex linguistic practice and little is known about how English language learners construct arguments in English. We investigated how undergraduate science students' English language proficiency and history were associated with their level of reasoning in scientific arguments ($N = 166$). We found that participants' English language experiences, including their proficiency and history, were associated with the level of reasoning demonstrated. These findings suggest a need to identify the associated barriers faced by students from diverse language backgrounds and design equitable educational supports and assessments.

Summary: The sophistication of students' scientific arguments was correlated with their English language experience, suggesting a need to identify reasons for language barriers and design equitable supports for students whose first language is not English.

Introduction

Citizens need to be able to build arguments and reason with scientific evidence especially in a world facing complex socioscientific issues (*e.g.*, climate change, viral pandemics) (National Research Council, 2013; Organisation for Economic Cooperation and Development, 2006; Social Sciences and Humanities Research Council, 2018; United Nations, 2015). As many students learn scientific argumentation skills in English settings, the active language production required can underestimate the abilities of students from English-as-an-additional language backgrounds (Eng+) (Abedi, Zhang, Rowe, & Lee, 2020; Afitska & Heaton, 2019; C. Buxton et al., 2014; Curtis & Millar, 1988). 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). For example, 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). Educational researchers have analyzed written arguments to gain insight into students' reasoning and ultimately improve science education (Becker, Noyes, & Cooper, 2016; Bodé, Deng, & Flynn, 2019; Cooper, Kouyoumdjian, & Underwood, 2016; Deng & Flynn, 2021; A. Moon, Stanford, Cole, & Towns, 2016; Alena Moon, Moeller, Gere, & Shultz, 2019; 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 responses constructed in a single language (usually English), without taking potential language barriers into account.

Given language's potential barrier to learning an important societal skill and the importance of creating equitable, diverse, and inclusive educational environments (Abedi, 2002; Afitska & Heaton, 2019; C. A. Buxton, Salinas, Mahotiere, Lee, & Secada, 2015; Chin, 2010; Maerten-Rivera et al., 2010; National Academies of Science Engineering and Medicine, 2018; Pyburn, Pazicni, Benassi, & Tappin, 2013; Shaw, Bunch, & Geaney, 2010; Solano-Flores & Trumbull, 2003; Sotelo-Dynega, Ortiz, Flanagan, & Chaplin, 2013; Wolf et al., 2008; Yore et al., 2004), language's role needs to be better understood in educational contexts such as scientific argumentation (Lee & Stephens, 2020; National Academies of Science Engineering and Medicine, 2018). 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). For example, in the United States, over 4.6 million Eng+ students make up 9.4% of all classrooms (National Center for Education Statistics, 2017). Globally, the population of Eng+ students is rising. In the United Kingdom, the Eng+ student population has increased 3.7% since 2012 (21.2% of all students) (Oxley & de Cat, 2019); in the US, 38% of STEM graduate students are temporary visa holders from non-English nations, a number that has steadily increased since 2000 (National Science Foundation, 2020).

In this study, we investigated the potential association between English language proficiency and history with the sophistication of undergraduate postsecondary students' written arguments in chemistry. Previous work on language and learning has been conducted in K–12 mathematics, biology, or general science contexts. Few investigations have been conducted in chemistry (Pyburn et al., 2013) and post-secondary contexts (LaCosse, Canning, Bowman,

Murphy, & Logel, 2020), despite reports that chemistry's dependence on English-based nomenclature, models, and symbols may pose unique challenges for Eng+ learners (Childs, Markic, & Ryan, 2015).

Research questions

Our research questions were:

1. What differences in mode of reasoning might there in arguments constructed by postsecondary science students who have 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 mode of reasoning postsecondary science students employ when constructing an argument in chemistry?

Methods

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. With a diverse student population, many students speak a language other than English (University of Ottawa, 2021). We asked professors teaching chemistry courses at the University to forward a recruitment text to students in their courses *via* email. The final sample 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 (**Fig. 1**) (Alena Moon et al., 2019). Participants were also asked 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.
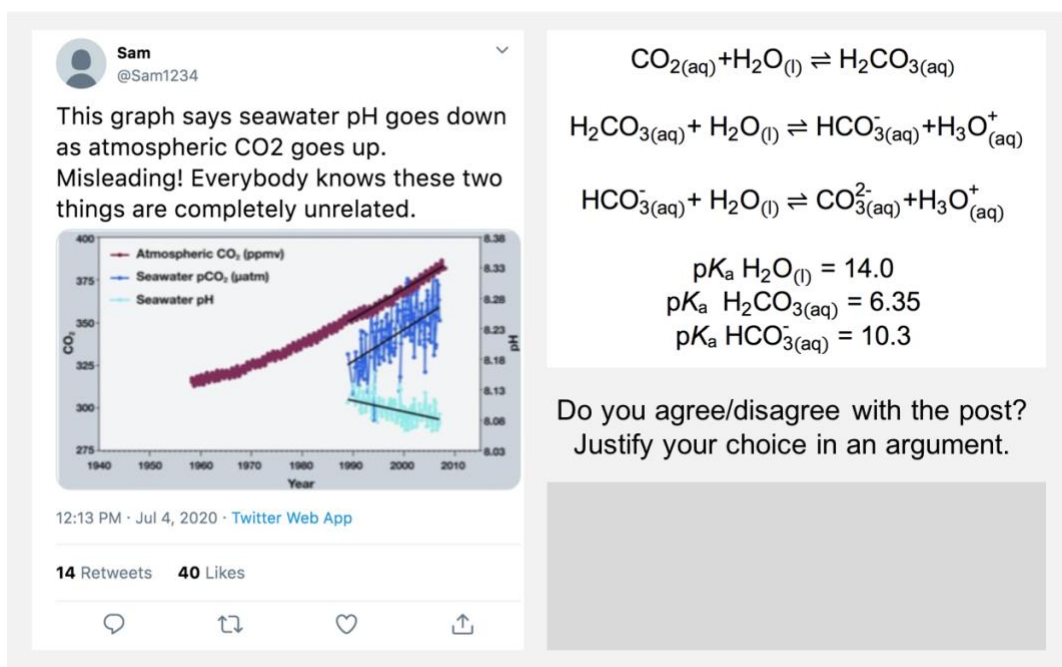
Fig. 1: Participants were asked to construct an argument to the social media post (left) using relevant chemical equilibria (right).

Standardized, open-ended interviews were conducted with a subset of participants (n = 15) to gain additional insight into participants' perceptions of how language contributes to learning and communicating in science.

Analysis

We analyzed each argument to identify the claim, evidence used, and mode of reasoning: descriptive, relational, linear causal, or multicomponent causal (Table 1) (Bodé et al., 2019; Carle, El Issa, Pilote, & Flynn, 2020; Caspari, Kranz, & Graulich, 2018; Deng & Flynn, 2021; A. Moon et al., 2016; Alena Moon, Stanford, Cole, & Towns, 2016; Moreira, Marzabal, & Talanquer, 2019; Sevian & Talanquer, 2014; Toulmin, 1958; Weinrich & Talanquer, 2016). *Descriptive* arguments list or give features and/or the properties of entities (*e.g.*, the 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.

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

Mode of reasoning	Definitions, adapted to the prompt	Examples
Descriptive	Argument describes explicit features of the graph or equations (*e.g.*, as atmospheric CO_2 concentration goes up, pH goes down).	Based on the equations, H_3O^+ is generated. Based on the graph, CO_2 goes up as pH goes down.

Relational	Argument identifies that increased CO_2 concentrations (atmospheric or aqueous) are correlated with H^+ being generated, using the equations. Identifies that increased H^+ is correlated with decreasing seawater pH.	Based on the graph, CO_2 goes up as pH goes down. As is shown in the equations, increasing CO_2 produces more H_3O^+ ions. Based on the equations, H_3O^+ is generated. Increasing H_3O^+ makes the seawater more acidic, so pH goes down.
Linear Causal	Argument uses Le Chatelier's principal or pK_a values to justify why atmospheric CO_2 concentrations influence the directions of the chemical equilibria and connects back to the impact that has on seawater pH.	Based on the graph, CO_2 goes up as pH goes down. Due to Le Chatelier's principle, increasing CO_2 produces carbonic acid, which leads to increased concentration of H_3O^+ ions. This results in increased acidification of the seawater, and thus a lower pH.

To generate language profiles for each participant, we created English language proficiency and history indices for each respondent using the LEAP-Q data (Figure S1) (Krizman, Skoe, Marian, & Kraus, 2014; Marian, Chabal, Bartolotti, Bradley, & Hernandez, 2014; Reichle & Birdsong, 2014).
5 Statistical tests were used to evaluate how participants' language backgrounds correlated with their demonstrated mode of reasoning. All statistical tests were conducted in SPSS (details in 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
10 variables: relative English language history and relative English language proficiency. 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 (categorial predictor). A second logistic regression was conducted with mode of reasoning as the dependent variable
15 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' arguments (descriptive, relational, causal).
20

Interviews were analyzed using educational research approaches (Cresswell, 2012; Lincoln & Guba, 1985), with particular attention to: (a) how participants perceived the role of language in their *own* learning experiences and (b) how they perceived the role of language in *others'* learning experiences (peers, friends, classmates).
25

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, χ^2 (1, N = 166) = 11.09, $p < 0.01$, $\phi = 0.253$ (Fig.
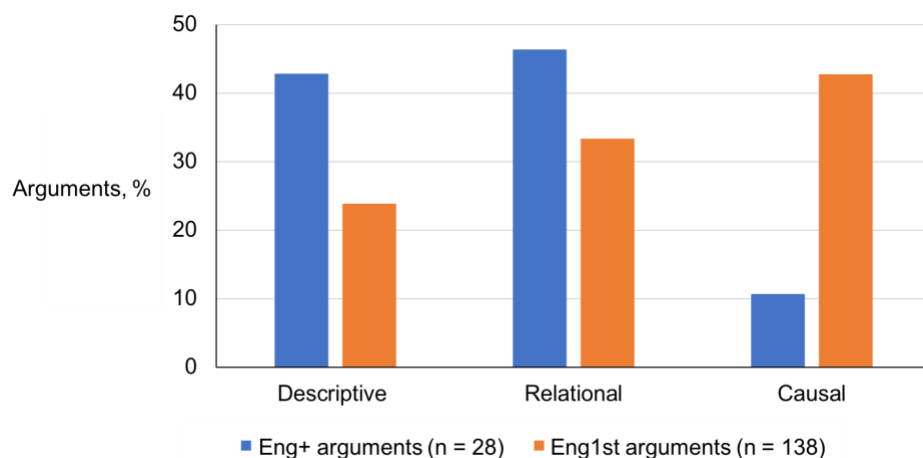30 2).

Fig. 2: Reasoning exhibited in scientific arguments. Eng+ (n = 28, blue, left) and Eng1st (n = 138, orange, right).

Participants with lower English history and proficiency scores produced the fewest causal arguments

LEAP-Q data analysis revealed heterogeneity in the Eng+ and Eng1st samples, which has been cited as a common challenge when categorizing individuals into language groups (Lee & Stephens, 2020; National Academies of Science Engineering and Medicine, 2018). For example, participants who were relatively new to English or experienced with English as an additional language are both considered Eng+. Therefore, to build categorizations from the data, we used cluster analysis to build groupings based on participants' LEAP-Q data. This analysis uncovered four language clusters (**Fig. 3**).
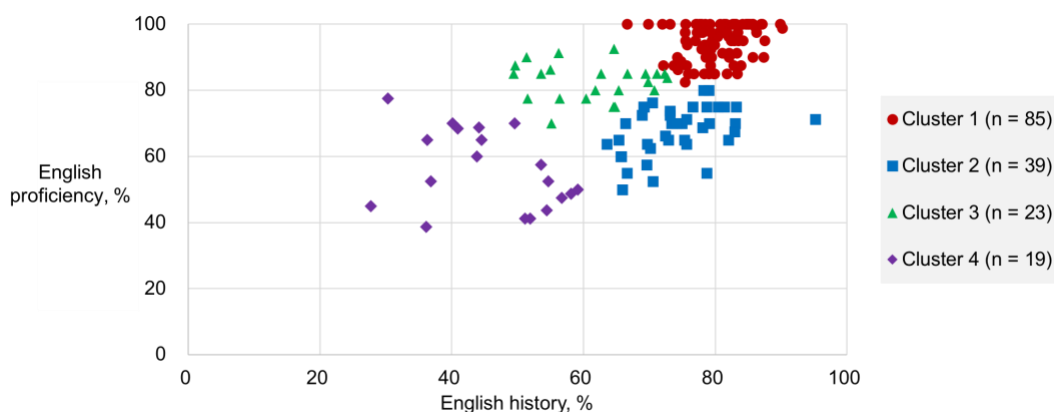

Fig. 3: Participants' English proficiency and history, organized into four clusters using cluster analysis (N = 166).

Cluster 1 responses more frequently exhibited causal reasoning compared to Cluster 4, χ^2 (6, N = 166) = 14.02, $p < 0.05$, $\phi = 0.284$ (**Fig. 4**). Cluster 1 had the highest proportion of participants who identified as Eng1st (98.8%), χ^2 (9, N = 166) = 94.97, $p < 0.0001$, $\phi = 0.756$ (**Fig. 4**). Of those, 91% having 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.

Fig. 4a: Language clusters vs. modes of reasoning   **Fig. 4b: Language clusters vs. primary general/chemistry language**
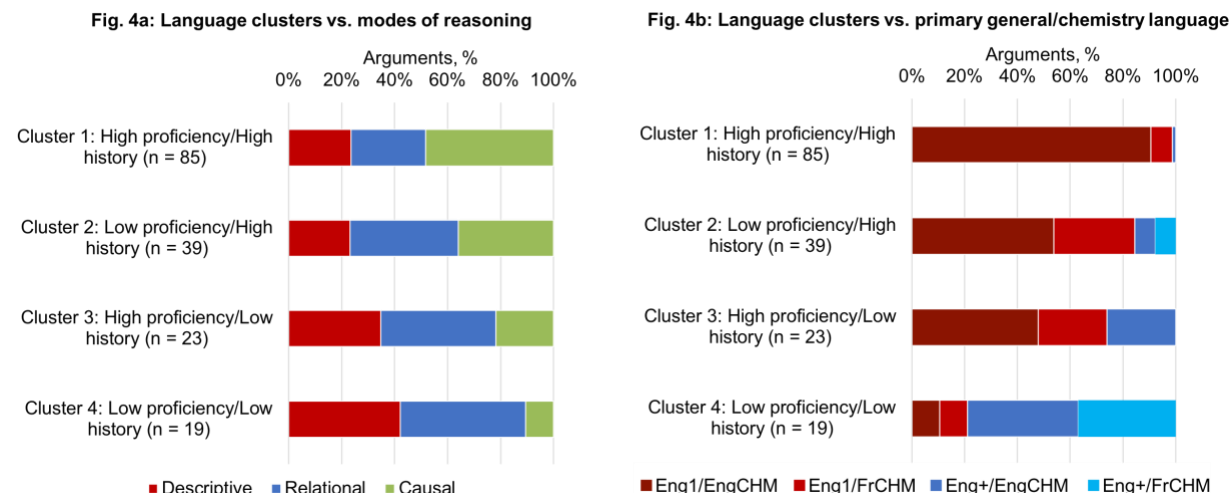

Fig. 4a: Distributions of the modes of reasoning for each language cluster; Fig. 4b: Distributions of general/chemistry language backgrounds for each language cluster.

Higher English history was associated with a greater proportion of causal arguments

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 S2). 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, Exp(B) = 1.085, $p < 0.05$ (Table S3).

Eng+ participants who had learned chemistry primarily in French did not demonstrate causal reasoning at all in their arguments (Fig. 5). 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). We note that there may be other variables in play here beyond language—for example, perhaps English and French chemistry courses expect different types of reasoning from students.
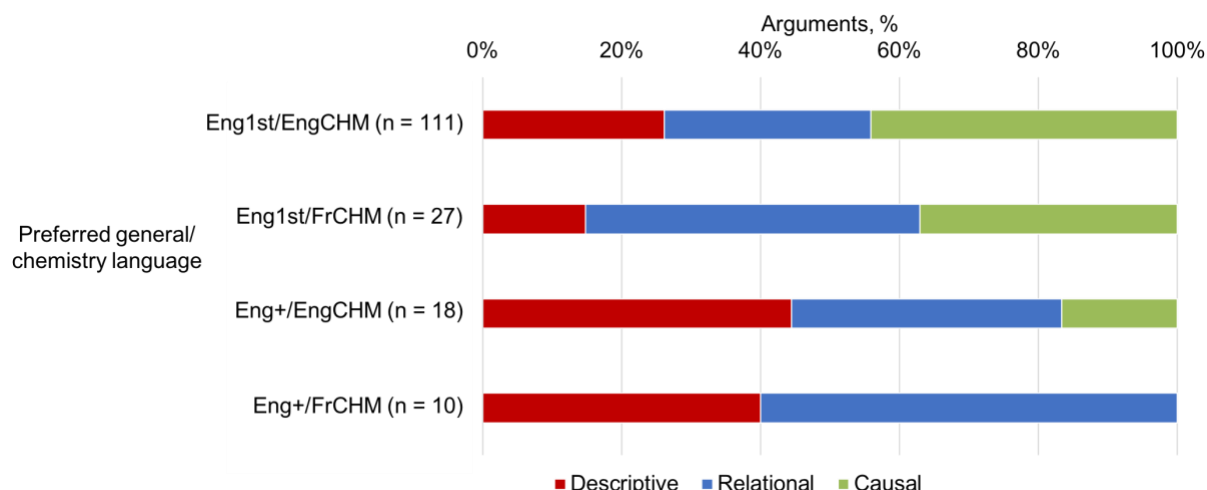
Fig. 5: Intersection between language group (Eng+, Eng1st), primary language of chemistry instruction (EngCHM, FrCHM) and reasoning.

Language of preference for communicating chemistry was connected with general and chemistry-specific language experiences

Across the interview sample, participants' preferred language for communicating chemistry often reflected the language they had learned chemistry in, regardless of their first language (Fig. 6, with examples in Table S7). Both Eng+ and Eng1st participants who had learned the majority of their chemistry in English said they preferred English in chemistry contexts and it would be difficult for them to communicate the same ideas in their first and/or home languages. In contrast, Eng+ participants learning chemistry in French stated greater comfort with French than English; they believed French allowed them to communicate more effectively and provide more detail than English.

Several Eng1st and Eng+ participants who had learned chemistry in English said they could also communicate chemistry effectively in French given the bilingual context; however, these participants noted they would mentally translate from English to French and preferred communicating chemistry ideas in English. Other Eng1st participants learning chemistry in French said they often could think about chemistry in both French and English and would be able to just as easily generate a response in French as in English.

Several participants who had mixed language experiences in chemistry (e.g., both French and English) added that their preference for communicating *specific* ideas could also depend on the language in which those ideas were taught. For example, if the student was taught their high school chemistry in French but university-level chemistry in English, they would prefer to communicate the high school ideas in French and the university-level ideas in English.
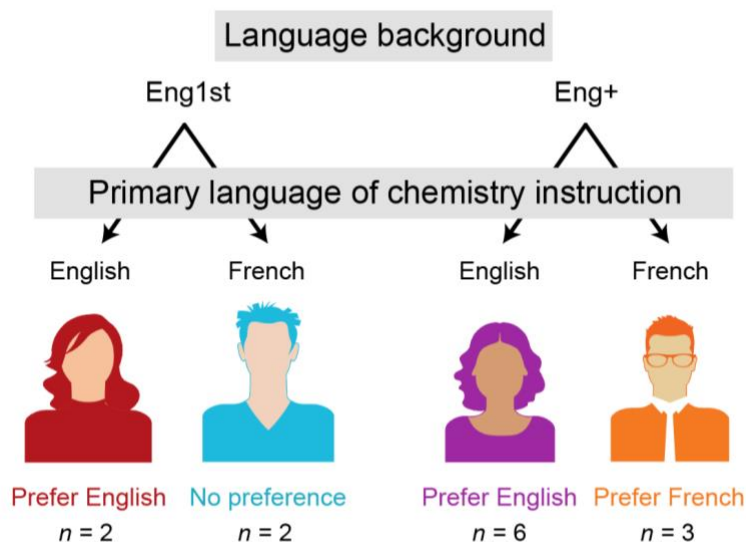
Fig. 6: Based on interview data, participants' comfort with responding in English vs. French was connected both their general language background (Eng1st or Eng+) and their primary language of chemistry instruction (English or French).

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' preferred language for communicating chemistry often reflected the language they had learned chemistry in.

Taken together, the findings suggest that language ability is correlated with students' performance on written assessments in chemistry. Importantly, communicating in English is challenging for students from both non-English home and school backgrounds. This finding may have broader implications for equity within science. 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 context, consequently limiting their participation in science or making participation much more challenging (Cheng et al., 2019; Elnathan, 2021; Ortega, 2021). For example, non-English scientists 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, although during recruitment, 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 (Liamputtong, 2008). Future endeavors 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). 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 S6). 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; Lee, 2005).

Conclusions

Constructing scientific arguments is an essential skill for members of society, especially in a world facing complex socioscientific issues. Scientific argumentation is a complex linguistic practice but there is limited research on how individuals from diverse language backgrounds construct those arguments. Therefore, we investigated how undergraduate science students' English language proficiency and history were associated with their level of reasoning in scientific arguments, using an online data collection instrument that asked student participants to analyze data related to climate change.

We found that participants' English language experiences, including their proficiency and history, were associated with the level 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). English language history was a significant predictor of the reasoning exhibited in participants' arguments.

The language of science instruction seemed to play an important role, as some participants described a greater 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. Language background was associated with the reasoning demonstrated in participants' arguments; no participant from an Eng+ background who had learned chemistry primarily in French demonstrated causal reasoning in their responses. The lack of causal responses could be due to language barriers or previous expectations in French courses where causal reasoning was not taught or assessed; more research is needed.

In an educational setting such as exams, the responses of Eng+ students may not accurately reflect their knowledge and reasoning capabilities. Language barriers could also hinder knowledge and skill acquisition. In later career settings, Eng+ professionals who have learned chemistry in a non-English language 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 determine whether the same findings are demonstrated in higher stakes settings (e.g., exams) and to 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), training for educators (e.g., designing inclusive assessments, mitigating implicit bias when evaluating work), and supports for Eng+ learners (e.g., translation software).

Creating a more equitable and inclusive society requires mitigating barriers that hinder participation of individuals from specific cultures and experiences, including individuals from diverse language backgrounds. By better understanding language-related barriers to learning in science (including on assessment), we can better support students and scientists throughout their careers, enhancing diversity and innovation in science (Bell, Villado, Lukasik, Belau, & Briggs, 2011; Harris, Mack, Bryant, Theobald, & Freeman, 2020; Hofstra et al., 2020; Impellizzeri & Coe, 2021; Lee, Quinn, & Valdés, 2013; McKinsey & Company, 2015; Valantine, Collins, & Verma, 2015).

References and Notes

Abedi, J. (2002). Assessment and accommodations of English language: Issues, concerns, and recommendations. *Journal of School Improvement*, *3*(1), 83–89.

Abedi, J., Zhang, Y., Rowe, S. E., & Lee, H. (2020). Examining Effectiveness and Validity of Accommodations for English Language Learners in Mathematics: An Evidence-Based Computer Accommodation Decision System. *Educational Measurement: Issues and Practice*, *39*(4), 41–52. https://doi.org/10.1111/emip.12328

Afitska, O., & Heaton, T. J. (2019). Mitigating the effect of language in the assessment of science: A study of English-language learners in primary classrooms in the United Kingdom. *Science Education*, 1396–1422. https://doi.org/10.1002/sce.21545

Amano, T., González-Varo, J. P., & Sutherland, W. J. (2016). Languages Are Still a Major Barrier to Global Science. *PLoS Biology*, *14*(12), e2000933. https://doi.org/10.1371/journal.pbio.2000933

Becker, N., Noyes, K., & Cooper, M. M. (2016). Characterizing Students' Mechanistic Reasoning about London Dispersion Forces. *Journal of Chemical Education*, *93*(10), 1713–1724. https://doi.org/10.1021/acs.jchemed.6b00298

Bell, S. T., Villado, A. J., Lukasik, M. A., Belau, L., & Briggs, A. L. (2011). Getting specific about demographic diversity variable and team performance relationships: A meta-analysis. *Journal of Management*, *37*(3), 709–743. https://doi.org/10.1177/0149206310365001

Bodé, N. E., Deng, J. M., & Flynn, A. B. (2019). Getting Past the Rules and to the WHY: Causal Mechanistic Arguments When Judging the Plausibility of Organic Reaction Mechanisms. *Journal of Chemical Education*, *96*(6), 1068–1082. https://doi.org/10.1021/acs.jchemed.8b00719

Buxton, C. A., Salinas, A., Mahotiere, M., Lee, O., & Secada, W. (2015). Fourth-grade emergent bilingual learners' scientific reasoning complexity, controlled experiment practices, and content knowledge when discussing school, home, and play contexts. *Teachers College Record*, *117*(2).

Buxton, C., Allexsaht-Snider, M., Aghasaleh, R., Kayumova, S., Kim, S.-H., Choi, Y.-J., & Cohen, A. (2014). Potential Benefits of Bilingual Constructed Response Science Assessments for Understanding Bilingual Learners' Emergent Use of Language of Scientific Investigation Practices. *Double Helix*, *2*(September 2016), 1–21.

Carle, M. S., El Issa, R., Pilote, N., & Flynn, A. B. (2020). Ten essential delocalization learning outcomes: How well are they achieved? *ChemRxiv*, 1–28.

Caspari, I., Kranz, D., & Graulich, N. (2018). Resolving the complexity of organic chemistry students ' reasoning through the lens of a mechanistic framework. *Chem. Educ. Res. Pract.*, *19*, 1117–1141. https://doi.org/10.1039/C8RP00131F

Cheng, Y., Dharwadkar, S., Sheridan, V., Rojas, C. R., Bosch Grau, M., & Gordin, M. (2019). Science's language barrier. *Nature*, *570*, 265.

Childs, P. E., Markic, S., & Ryan, M. C. (2015). The Role of Language in Teaching and Learning Chemistry. In J. Garcia-Martinez & E. Serrano-Torregrossa (Eds.), *Chemistry Education: Best Practices, Opportunities and Trends* (pp. 421–445). Wiley-VCH Verlag GmbH & Co. KGaA.

Chin, W. Y. (2010). Linguistic profiling in education: how accent bias denis equal educational opportunites to students of color. *The Scholar*.

Cooper, M. M., Kouyoumdjian, H., & Underwood, S. M. (2016). Investigating Students' Reasoning about Acid–Base Reactions. *Journal of Chemical Education*, *93*(10), 1703–1712. https://doi.org/10.1021/acs.jchemed.6b00417

Cresswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Boston, MA: Pearson.

Curtis, S., & Millar, R. (1988). Language and Conceptual Understanding in Science: A comparison of English and Asian language speaking children. *Research in Science & Technological Education*, *6*(1), 61–77. https://doi.org/10.1080/0263514880060106

Deng, J. M., & Flynn, A. B. (2021). Reasoning, granularity, and comparisons in students' arguments on two organic chemistry items. *Chemistry Education Research and Practice*, *22*, 749–771. https://doi.org/10.1039/x0xx00000x

Elnathan, R. (2021). English is the language of science — but precision is tough as a non-native

speaker. Retrieved from https://www.nature.com/articles/d41586-021-00899-y

González-Howard, M., & Suárez, E. (2021). Retiring the term English language learners: Moving toward linguistic justice through asset-oriented framing. *Journal of Research in Science Teaching*, (January), 749–752. https://doi.org/10.1002/tea.21684

Hanauer, D. I., & Englander, K. (2011). Quantifying the Burden of writing research articles in a second language: Data from Mexican scientists. *Written Communication*, *28*(4), 403–416. https://doi.org/10.1177/0741088311420056

Hanauer, D. I., Sheridan, C. L., & Englander, K. (2019). Linguistic Injustice in the Writing of Research Articles in English as a Second Language: Data From Taiwanese and Mexican Researchers. *Written Communication*, *36*(1), 136–154. https://doi.org/10.1177/0741088318804821

Harris, R. B., Mack, M. R., Bryant, J., Theobald, E. J., & Freeman, S. (2020). Reducing achievement gaps in undergraduate general chemistry could lift underrepresented students into a "hyperpersistent zone." *Science Advances*, *6*(24), 1–9. https://doi.org/10.1126/sciadv.aaz5687

Hofstra, B., Kulkarni, V. V., Galvez, S. M. N., He, B., Jurafsky, D., & McFarland, D. A. (2020). The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(17), 9284–9291. https://doi.org/10.1073/pnas.1915378117

Huang, J. (2008). How accurate are ESL students' holistic writing scores on large-scale assessments?-A generalizability theory approach. *Assessing Writing*, *13*(3), 201–218. https://doi.org/10.1016/j.asw.2008.10.002

Impellizzeri, S., & Coe, I. R. (2021). The complex chemistry of diversity and inclusion: a 30-year synthesis. *Canadian Journal of Chemistry*, *99*(8), 653–660. https://doi.org/10.1139/cjc-2021-0063

Krizman, J., Skoe, E., Marian, V., & Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain and Language*, *128*(1), 34–40. https://doi.org/10.1038/jid.2014.371

LaCosse, J., Canning, E. A., Bowman, N. A., Murphy, M. C., & Logel, C. (2020). A social-belonging intervention improves STEM outcomes for students who speak English as a second language. *Science Advances*, *6*(40), 1–11. https://doi.org/10.1126/sciadv.abb6543

Lee, O. (2005). Science education with English language learners: Synthesis and research agenda. *Review of Educational Research*, *75*(4), 491–530. https://doi.org/10.3102/00346543075004491

Lee, O., Quinn, H., & Valdés, G. (2013). Science and Language for English Language Learners in Relation to Next Generation Science Standards and with Implications for Common Core State Standards for English Language Arts and Mathematics. *Educational Researcher*, *42*(4), 223–233. https://doi.org/10.3102/0013189X13480524

Lee, O., & Stephens, A. (2020). English Learners in STEM Subjects: Contemporary Views on

STEM Subjects and Language With English Learners. *Educational Researcher*, *49*(6), 426–432. https://doi.org/10.3102/0013189X20923708

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

Liamputtong, P. (2008). *Doing Cross-Cultural Research: Ethical and Methodological Perspectives*.

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage Publications Inc.

Lindsey, P., & Crusan, D. (2011). How Faculty Attitudes and Expectations Toward Student Nationality Affect Writing Assessment. *Across the Disciplines*, *8*(4), 1–19. https://doi.org/10.37514/atd-j.2011.8.4.23

Lyon, E. G., Bunch, G. C., & Shaw, J. M. (2012). Navigating the language demands of an inquiry-based science performance assessment: Classroom challenges and opportunities for English learners. *Science Education*, *96*(4), 631–651. https://doi.org/10.1002/sce.21008

Maerten-Rivera, J., Myers, N., Lee, O., & Penfield, R. (2010). Student and school predictors of high-stakes assessment in science. *Science Education*, *94*(6), 937–962. https://doi.org/10.1002/sce.20408

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Marian, V., Chabal, S., Bartolotti, J., Bradley, K., & Hernandez, A. E. (2014). Differential recruitment of executive control regions during phonological competition in monolinguals and bilinguals. *Brain and Language*, *139*, 108–117. https://doi.org/10.1016/j.bandl.2014.10.005

McKinsey & Company. (2015). Why diversity matters. Retrieved September 3, 2020, from https://www.mckinsey.com/business-functions/organization/our-insights/why-diversity-matters#

Milnes, T., & Cheng, L. (2008). Teachers' Assessment of ESL Students in Mainstream Classes: Challenges, Strategies, and Decision-Making. *TESL Canada Journal*, *26*(1), 49. https://doi.org/10.18806/tesl.v26i1.129

Moon, A., Stanford, C., Cole, R., & Towns, M. (2016). The nature of students' chemical reasoning employed in scientific argumentation in physical chemistry. *Chemistry Education Research and Practice*, *17*(2), 353–364. https://doi.org/10.1039/c5rp00207a

Moon, Alena, Moeller, R., Gere, A. R., & Shultz, G. V. (2019). Application and testing of a framework for characterizing the quality of scientific reasoning in chemistry students' writing on ocean acidification. *Chemistry Education Research and Practice*, *20*(3), 484–494. https://doi.org/10.1039/c9rp00005d

Moon, Alena, Stanford, C., Cole, R., & Towns, M. H. (2016). The nature of students' chemical reasoning employed in scientific argumentation in physical chemistry. *Chemistry Education Research and Practice*, *17*(2), 353–364. https://doi.org/10.1039/c5rp00207a

Moreira, P., Marzabal, A., & Talanquer, V. (2019). Using a mechanistic framework to characterise chemistry students' reasoning in written explanations. *Chemistry Education Research and Practice*, *20*(1), 120–131. https://doi.org/10.1039/c8rp00159f

National Academies of Science Engineering and Medicine. (2018). *English Learners in STEM Subjects: Transforming Classrooms, Schools, and Lives*. Washington, D.C.: The National Academies Press. https://doi.org/10.17226/25182

National Center for Education Statistics. (2017). *The Condition of Education: English Language Learners in Public Schools*.

National Research Council. (2013). *Next Generation Science Standards*. Press. Retrieved from https://www.nextgenscience.org/

National Science Foundation. (2020). *2019 Doctorate Recipients from U.S. Universities* (Vol. NSF 21-308). Retrieved from http://www.nsf.gov/statistics/2016/nsf16300/digest/nsf16300.pdf

Organisation for Economic Cooperation and Development. (2006). *Assessing scientific, reading and mathematical literacy: a framework for PISA 2006*.

Ortega, R. P. (2021). Science's English dominance hinders diversity—but the community can work toward change. *Science*. Retrieved from https://www.sciencemag.org/careers/2020/10/science-s-english-dominance-hinders-diversity-community-can-work-toward-change

Osborne, J. F., & Patterson, A. (2011). Scientific Argument and Explanation: A Necessary Distinction? *Science Education*, *95*(4), 627–638. https://doi.org/10.1002/sce.20438

Oxley, E., & de Cat, C. (2019). A systematic review of language and literacy interventions in children and adolescents with English as an additional language (EAL). *Language Learning Journal*, *0*(0), 1–23. https://doi.org/10.1080/09571736.2019.1597146

Project Atlas. (2019). *Research and Insights: Global Mobility Trends*. Retrieved from https://www.iie.org/Research-and-Insights/Project-Atlas/Explore-Data/Infographics/2019-Project-Atlas-Infographics

Pyburn, D. T., Pazicni, S., Benassi, V. A., & Tappin, E. E. (2013). Assessing the relation between language comprehension and performance in general chemistry. *Chemistry Education Research and Practice*, *14*(4), 524–541. https://doi.org/10.1039/c3rp00014a

Ramírez-Castañeda, V. (2020). Disadvantages in preparing and publishing scientific papers caused by the dominance of the English language in science: The case of Colombian researchers in biological sciences. *PLoS ONE*, *15*(9 September). https://doi.org/10.1371/journal.pone.0238372

Reichle, R. V., & Birdsong, D. (2014). Processing focus structure in L1 and L2 French : L2 proficiency effects on ERPs. *Studies in Second Language Acquisition*, *36*(3), 535–564. https://doi.org/10.1017/S0272263113000594

Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing Reasoning in Student Scientific Inquiry: A Framework for Discourse Analysis Developed From Philosophy of

Science. *Science Education*, *92*(3), 499–525. https://doi.org/10.1002/sce.20264

Sevian, H., & Talanquer, V. (2014). Rethinking chemistry: a learning progression on chemical thinking. *Chem. Educ. Res. Pract.*, *15*(1), 10–23. https://doi.org/10.1039/c3rp00111c

Shaw, J. M., Bunch, G. C., & Geaney, E. R. (2010). Analyzing language demands facing English learners on science performance assessments: The SALD framework. *Journal of Research in Science Teaching*, *47*(8), 909–928. https://doi.org/10.1002/tea.20364

Social Sciences and Humanities Research Council. (2018). Truth Under Fire in a Post-Fact World. Retrieved September 1, 2019, from https://horizons.gc.ca/en/2018/10/19/the-next-generation-of-emerging-global-challenges/#truth-under-fire

Solano-Flores, G., & Trumbull, E. (2003). Examining Language in Context: The Need for New Research and Practice Paradigms in the Testing of English-Language Learners. *Educational Researcher*, *32*(2), 3–13. https://doi.org/10.3102/0013189X032002003

Sotelo-Dynega, M., Ortiz, S. O., Flanagan, D. P., & Chaplin, W. F. (2013). English language proficiency and test performance: an evaluation of bilingual students with the Woodcock-Johnson III Tests of Cognitive Abilities. *Psychology in the Schools*, *50*(8), 781–797. https://doi.org/10.1002/pits

Swanson, L. H., Bianchini, J. A., & Lee, J. S. (2014). Engaging in argument and communicating information: A case study of english language learners and their science teacher in an urban high school. *Journal of Research in Science Teaching*, *51*(1), 31–64. https://doi.org/10.1002/tea.21124

Toulmin, S. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press. https://doi.org/10.1080/00048405985200191

United Nations. (2015). Transforming our World: the 2030 Agenda for Sustainable Development. Retrieved from https://www.un.org/sustainabledevelopment/sustainable-development-goals/

University of Ottawa. (2021). Francophonie and Bilingualism. Retrieved October 7, 2021, from https://www.uottawa.ca/francophonie-bilingualism/

Valantine, H. A., Collins, F. S., & Verma, I. M. (2015). National Institutes of Health addresses the science of diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(40), 12240–12242. https://doi.org/10.1073/pnas.1515612112

Watts, F. M., Schmidt-Mccormack, J. A., Wilhelm, C. A., Karlin, A., Sattar, A., Thompson, B. C., … Shultz, G. V. (2020). What students write about when students write about mechanisms: Analysis of features present in students' written descriptions of an organic reaction mechanism. *Chemistry Education Research and Practice*, *21*(4), 1148–1172. https://doi.org/10.1039/c9rp00185a

Weinrich, M. L., & Talanquer, V. (2016). Mapping students' modes of reasoning when thinking about chemical reactions used to make a desired product. *Chemistry Education Research and Practice*, *17*(2), 394–406. https://doi.org/10.1039/c5rp00208g

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

(2008). Issues in assessing English language learners: English language proficiency measures and accommodation uses (CRESST Report 732). *Review Literature And Arts Of The Americas*, 1–96.

Yore, L. D., Hand, B., Goldman, S. R., Hildebrand, G. M., Osborne, J. F., Treagust, D. F., & Wallace, C. S. (2004). New directions in language and science education research. *Reading Research Quarterly*, *39*(3), 347–352. https://doi.org/10.1598/RRQ.39.3.8

Author contributions:

Conceptualization: JMD, ABF

Methodology: JMD, ABF

Investigation: JMD, MR

Visualization: JMD, MR

Funding acquisition: JMD, ABF

Project administration: JMD, ABF

Supervision: ABF

Writing – original draft: JMD

Writing – review & editing: JMD, ABF

Competing interests: Authors declare they have no competing interest.

Supporting Materials

Methods and additional materials

Supplemental text

Figs. S1 to S2

Tables S1 to S9

References (1-65)