The Effect of Socioscientific Topics on Discourse within an Online Game Designed to Engage Middle School Students in Scientific Argumentation

Jana Craig-Hare, Marilyn Ault, Amber Rowland
University of Kansas

To cite this article:

This article may be used for research, teaching, and private study purposes.

Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles.

The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material.
The Effect of Socioscientific Topics on Discourse within an Online Game Designed to Engage Middle School Students in Scientific Argumentation

Jana Craig-Hare, Marilyn Ault, Amber Rowland

Abstract

The purpose of this study was to investigate the types of argumentation discourse displayed by students when they engaged in chat as part of an online multiplayer game about both socioscientific and scientific topics. Specifically, this study analyzed discourse episodes created by middle school students as they discussed scientific and socioscientific topics within an online, multiplayer game. Using a Discourse Analysis Scoring Guide, student discussions were coded based on the type of interaction or statements made. Analysis included a comparison between the types of topics (scientific vs. socioscientific) and the student author’s justification for their decision to accept, reject or withhold judgement about the claim; teammate comments related to the author’s justification; an overall rating of the discourse episode interaction; and frequency of argumentation vocabulary use throughout the discourse episode. Results indicated that socioscientific topics produced collaborative discourse episodes that were positive, supportive, and civil within an argumentation framework.

Introduction

The ability to engage in productive discourse is a skill that has been recognized as key to learning. The theoretical perspective that learning can be socially constructed through conversation or discourse is well founded (Berland & Reiser, 2011; Clark & Sampson, 2007; Pellegrino & Hilton, 2013; Prestridge, 2009; Vygotsky, 1978). Argumentation is a type of discourse that involves a group of equal participants, or learners, engaged in the social construction of knowledge by specifically addressing evidence and reasoning to consider or advance a claim (Duschl & Osborne, 2002; Osborne et al., 2013; Toulmin, 2003; Toulmin, Rieke, & Janik, 1984). Argumentation is a cross-curricular skill that is difficult to teach. Research suggests that integrating argumentation into science instruction is a significant challenge, both for teachers and students (Alozie, Moje, & Krajcik, 2010; Bulgren & Ellis, 2010). Because of this difficulty, a game has been developed that can be used to engage students in learning the knowledge and skills related to argumentation. This study addressed whether the chat section of the online game environment can engage students in quality argumentation discourse with either scientific or socioscientific content.

The need for students to develop skills and knowledge related to argumentation is reflected in both the Common Core State Standards (CCSS) (2010) and the Next Generation Science Standards (NGSS) (Achieve, 2013). The CCSS reflect an integrated view of reading, writing, speaking/listening, and argumentation across content areas, including science, mathematics, social studies, and English language arts. Argumentation skills encourage thoughtful student discourse by creating an environment in which students question each other’s claims and evaluate the strength of their evidence. These standards emphasize the need for students to know how to take a more critical stance when confronted with an argument; evaluate the quality of what they read, see, or hear; and defend their claims with appropriate evidence and reasoning. It has been noted that argumentation skills help increase students’ achievement and content knowledge by requiring them to think deeply about content, construct their own understanding of content, and apply it as they construct their arguments or critique those of others (Berland & Reiser, 2011; Clark & Sampson, 2007; Nussbaum, 2008; Osborne, Simon, Christodoulou, Howell-Richardson, & Richardson, 2013; Pellegriino & Hilton, 2013, Scheuer, Loll, Pinkwart, & McLaren, 2010).

The critical role of discourse, particularly argumentation, in students’ understanding of and learning about science has also become increasingly evident. The NRC (2012) notes that “science is fundamentally a social enterprise” (p. 27) where scientists engage in ongoing discourse with their colleagues, informally and formally, to share insights, brainstorm, and problem-solve. As defined by the NGSS, argumentation is “a mode of logical discourse used to clarify the strength of relationships between ideas and evidence that may result in revision of
an explanation” (Achieve, 2013). Toulmin, Rieke, and Janik (1984) defined argumentation as “the whole activity of making claims, challenging them, backing them up by producing reasons, criticizing those reasons, rebutting those criticisms, and so on” (p. 14). Middle school standards now require students to engage in scientific argumentation as a critical science practice.

**Using Argumentation in Socioscientific Discourse**

When presented with open-ended, controversial issues, students are empowered to discuss science-related topics that shape their current world and have a large impact on their future (Driver, Newton, & Osborne, 2000; Kolstø, 2001). These dynamic interactions between science and society focus not only on the issues behind science, but also the relationship with social, political, economic, and moral challenges (Sadler & Fowler, 2006). Discourse about socioscientific topics involves the skills of identifying evidence, reasoning, evaluating information, and the development of conceptual understandings (Sadler, 2004). Argumentation is also an important part of decision-making (Patronis, Potari, & Spiliotopoulou, 1999) when dealing with socioscientific issues (Fleming, 1986; Kolstø, 2001; Zeidler, 2003). Practice in argumentation (Duschl & Osborne, 2002; Kuhn, 1993) in the context of controversial issues is needed for making informed decisions, which is considered vital for developing scientifically literate students and advancing democratic societies (Aikenhead, 1985; Fullinwider, 1987; Kolstø, 2001).

Socioscientific content-based scenarios address issues that are personally meaningful and engaging to students. Oftentimes, they are controversial in nature but the topics have an added element of requiring a degree of moral reasoning or the evaluation of ethical concerns that are personally relevant in the process of arriving at decisions regarding possible resolution of those issues (Zeidler & Nichols, 2009; Chang & Chiu, 2008). These topics mirror issues found in modern society and connect to student lives through their environment, media, and personal interests. Components of argumentation provide a structure for students to discuss these contentious topics in a productive and meaningful manner.

**Learning Argumentation in a Computer-Mediated Environment**

A number of reviews summarize the history of computer-supported learning and, specifically, computer-support of argumentation skill development (Scheuer, Loll, Pinkwart, & McLaren, 2010; Sollier, Martinez, Jermann, & Muehlenbrock, 2005). These results, as well as the work of Linn and her colleagues and others, have demonstrated the ability to engage students in discourse and argumentation in scaffolded and controlled web-based spaces (Jeong & Joung, 2007; Linn, Clark, & Slotta, 2003; Linn & Eylon, 2011). The review by Sollier et al. (2005) addressed computer-supported applications that were designed to support collaborative learning. They identified features that were characteristic of successful collaborative learning environments. While their interest was specifically in whether it is possible to design online learning environments to be facilitated by a coach, the framework they described informs online instructional environments, in general. The salient characteristics of the instructional environments that can be applied to supporting argumentation in an online environment include a shared work-space that supported a social awareness of teammates, a chat function allowing for open-ended interactions, delineated roles, problem-solving actions, and graphical visualizations of performance. The chat and graphic visualizations were intended to give students a metacognitive perspective of their discourse actions.

Scheuer, Loll, Pinkwart, and McLaren’s (2010) more recent work reviewed a collection of software applications that were successful in teaching students the components of scientific argumentation. They identified five different types of support for argumentation, including free-form arguments, arguments based on transcripts, and system-provided prompts and examples. They concluded that software could be designed and implemented to support the development of the complex skill of argumentation. They suggested that by scaffolding good argumentation practices, the systems not only supported students in “learning to argue” but also supported “arguing to learn,” helping students learn about specific domain topics through argumentation (pg. 45). The systems they reviewed, however, were client-based and ran on individual computers. As such, they were not networked and tended to be for single users (Groesser, Gernsbacher, & Goldman, 2012). The users, therefore, learned the components of argumentation but did not engage in the practice with peers.

The increased use of online environments, such as chat, forums, or blogs, provides an additional space in which to observe and quantify discourse outside of specifically structured applications. For example, Jepson (2005) developed a scoring protocol for discourse that occurs in text and voice chat rooms. In these unstructured environments, he identified two roles – speakers (initiators) and interlocutors (responders) – and could quantify
negotiation of meaning and feedback for both roles. Chen and Chiu (2008) looked at online discussions in college-based discussion forums (such as Blackboard). Even though they were looking at a small number of participants and number of posts overall, they could quantify the flow of discussion. They identified five different types of messages and described the message properties. The online message types included: evaluation, knowledge statement, social statement, personal information, and elicitation. The message properties included agreement, disagreement, unresponsive/new topic, contribution, repetition, null content, positive social cue, negative social cue, and non-personal social cue. Chen and Chiu were interested in the ongoing exchanges and if the types and properties of comments predicted further types of comments. They examined how the flow of a discussion predicted later messaging and demonstrated statement properties of disagreement, contribution, social cue, and past visits can affect the properties of subsequent messages.

Many researchers have specifically attempted to quantify and then describe argumentation in online environments. Clark and Sampson (2007) developed an analytic framework for assessing argumentation in these more open online science learning environments (Clark, Sampson, Weinberger & Erkens, 2007). Based on the previous work of Erduran, Osborne, and Simon (Erduran, Simon, & Osborne, 2004; Simon, Osborne, & Erduran, 2003) Clark and Sampson use a strategy to score what they identified as discourse moves, the use of evidence and reasoning, and the conceptual quality in asynchronous threaded discussions in online environments. Their analysis described the overall quality of the online argumentation discussion, with the purpose of the discussion being to reach an agreement. This is because, they argue, from scientists’ perspective, the role of argumentation is persuasion in the process of developing new knowledge, the definition of dialogic communication. They see argumentation as both a social and a collaborative process and recognize, therefore, that many statements made in the process of argumentation cannot be defined based solely on Toulmin’s model. They suggest that there are elements to online (as well as face-to-face) argumentation that involve other operations, such as requests for clarification or statements of support for another’s claim. Their analysis included eight different types of discourse moves, some specifically reflecting Toulmin’s model and others describing social interactions. These more social statements included items such as: changing a claim, providing clarification, providing support, asking a question about meaning, requesting clarification about meaning, as well as social organizational comments and social but off-task comments. Given the social nature of argumentation, these discourse elements provide an organizing and supportive role that facilitates continuing the conversation in an egalitarian manner. Lu, Chiu, and Law (2011) expanded Clark and Sampson’s coding protocol. Based on their analysis of online interactions they expanded the coding of elements addressing Toulmin’s model and introduced two additional types of disagreements: disagreement with added justifications and disagreement against earlier justification. Overall, they agreed with Clark and Sampson that episodes of argumentation in an online environment can occur and be quantified; that the cognitive and social communicative processes of argumentation are closely related in online argumentation, as they are in a face-to-face episode; and that online and face-to-face argumentation interactions differ in use of evidence and explanations. It is clear, therefore, that students can learn and engage in a robust dialogue during argumentation in online spaces. These actions can include complex discourse moves and negation of meaning and are consistent with the interactions that re desired when addressing socioscientific content.

### Argumentation Using Chat within a Multiplayer Online Game

Supporting argumentation using chat within a game combines the free form of the chat environment with the many motivational aspects of a game environment. Argumentation that is included as a part of a game has the potential to engage participants more than a course-based online chat or threaded discussion. This is because games can be specifically designed to include features that create a heightened emotional attachment during play, resulting in a level of engagement that does not occur with typical online or face-to-face instruction. Research on the effect of technology-based games has consistently shown positive results regarding motivation, persistence, curiosity, attention, and attitude toward learning (Shin, Sutherland, Norris, & Soloway, 2012). Early studies of online games demonstrate that many features are successful in engaging players. These include features such as social interaction, competition, and collaborative play (Malone, 1981), the social context of the game (Choi & Kim, 2004; Hsu & Lu, 2004), and competition (Koster, 2005).

One highly compelling feature is the opportunity for players to interact. Because of the rich and compelling environment, online games have been a place for discourse, and discourse analysis, since the late 1970s when the first multi-user games appeared (Brown & Bell, 2006; McEwan, Gutwin, Mandryk, & Nacke, 2012). Gee (1992, 1996, 1999) studied the components of discourse within the context of online games. He suggested that this type of online environment constitutes a rich space in which discourse emerges and allows for integrating language with the use of symbols and slang. Discourse that occurs within multi-player online games is
persistent, player-produced, and useful; generally focusing on problem-solving and sense-making. The multi-
user chat environments in games promote naturally occurring conversations between both known and
anonymous players. Not only are these interactions used to gather information about the game, they are also
used to instruct others and socialize about events in and outside the game (Brown & Bell, 2006; Ferrari,
Lessiter, & Freeman, 2011; Nardi, Ly, & Harris, 2007). As Steinkuehler (2006) suggests, the chat conversations
initially appear superficial because of the use of abbreviations, images, grammatical and spelling errors, and
slang. Her further analysis, consistent with the views of Gee and others, suggests that the chat conversations in
MMOGs (massively multiplayer online games) have the same level of complexity as off-line language. Because
of the suitability of the online chat environment to engage players, particularly youth, in discourse, a number of
researchers have recognized that chat, within a game environment, is a suitable space for the development of
scientific argumentation skills & discourse (Squire & Jan, 2007; Steinkuehler & Chmiel, 2006).

Research Focus

The purpose of this study, therefore, was to investigate the types of argumentation discourse displayed by
students when they engaged in chat as part of an online multiplayer game about both socioscientific and
scientific topics. The question was whether there was a difference in the discourse in an online environment
when the game content was socioscientific or scientific in nature. The game, Reason Racer, was used to present
the topics and engage the students in a game environment. The game engages students in the skills and
knowledge of scientific argumentation within a fast-paced, competitive game. The last part of the game provides
students with the opportunity to engage with their fellow players about the topic of game play in an unstructured
chat environment. Our interest is in the content and general characteristics of these discourse episodes during
chat, whether students applied their recently acquired skills in scientific argumentation as a part of the chat
conversation and if the nature of the topic (scientific vs. socioscientific) impacted student discourse.

Method

Participants and Setting

Over 500 middle school students from six school districts in the Midwest participated in the use of the Reason
Racer game during science instruction in the Fall of 2012. These schools were in both rural and suburban
districts. Students were enrolled in the 7th and 8th grade and were between 11 to 15 years old. Mixed-gender
classrooms were comprised of 49% female students, 46% male students, and 5% unreported; taught by seven
different teachers within the six schools. These students completed 937 discourse episodes through their Reason
Racer game-play sessions using multiple scenarios that were of both scientific and socioscientific nature.
Individual students completed at least one discourse episode, however some students completed as many as six
discourse episodes including introductory game play. The seven teachers volunteered to participate in the
project and use the Reason Racer game (described below) as part of their science instruction.

Procedures

This study utilized data generated from student game play during the Reason Racer game. Reason Racer is an
online multiplayer arcade-style game that contains four parts, each designed to engage players in skills and
knowledge related to scientific argumentation. When setting up play for students, the teacher assigns the game
by selecting from 40 different scenarios covering topics in physical science; life science; earth and space
science; and engineering, technology and the application of science. The different scenarios, selected and
developed to be interesting to middle school students, populate the content of the game’s challenges. Students
play the game with their peers by interacting with one game scenario. Other play sessions may use the same or
different scenarios. The areas of argumentation addressed in the game include understanding a claim, judging
evidence about a claim based on type (fact, opinion, data, or theory) and quality, determining the reasoning
(authority, theory, or logic), considering counterarguments and rebuttals, and making judgments, based on
Toulmin’s model (Toulmin, 2003; Toulmin, Rieke, & Janik, 1984). Students who played the game 10 times
across a two-month period as a part of instruction improved in every aspect of argumentation skill and judgment
and reported an increase in confidence and motivation to engage in science compared to students who did not
play (Ault, Craig-Hare, Frey, Ellis, & Bulgren, 2015).
The first part of *Reason Racer* orients the players to the game through a humorous 30-second video. This video is one of 40 possible short previews of specific elements of the game such as the content of a particular scenario, a review of a specific component of scientific argumentation, or advice about how to participate in the chat environment.

The second part of *Reason Racer* engages the players in a competitive, multiplayer rally-race game; alternating between challenges, or PitStops, and racing segments across a variety of racecourses. The PitStops require actions that are common to fast-paced games, such as matching, ranking, sorting, and discriminating, all within a competitive, rate-based game interface. Figure 1 shows six of the eight PitStops from one scenario as an example. This scenario presents the claim that a new technology for biofuel production could utilize an enzyme found in a panda’s digestive system to help convert plant matter to biofuel. The PitStops contain the content of the game, requiring students to identify components or make decisions about the claim, evidence, reasoning, and challenges to the claim. During game play, students attempt to move through each PitStop as quickly, and with as few errors, as possible. The competitive racing component, Figure 2, is completed between each PitStop. Students navigate various racing tracks’ turns and obstacles as quickly as possible to move to the next PitStop. The speed and accuracy of a player’s performance in the PitStop affect the speed with which his or her car can move through the next racing segment. Incorrect responses slow down the presentation of items in the PitStop, which discourages guessing. The experience of the racing component occurring between the challenges results in students completing the PitStops faster and more accurately (Ault, Craig-Hare & Frey, 2016) than with a no-race option in between PitStops.

Figure 1. A sample of reason racer pit-stop challenges

Figure 2. Reason racer staging area and sample racecourses
The third part of the game involves decision-making. Students read a brief article that reviews most of the content they just encountered. Their task is to decide whether to accept, reject, or withhold their decision about the claim and write a justification statement, as seen in Figure 3. This comment populates the final portion of the game, the discourse part, as seen in Figure 4. After making the decision and entering a statement supporting their decision about the claim, the players race to the end of the game and receive their scores. This score provides achievements that allow the player to be more competitive during the next round of play.

The fourth part of the game, seen in Figure 4, involves players interacting with the other players in a peer-scored discourse environment. This environment is an unstructured chat window, monitored by fellow students and the teacher. In this environment, each player in the game begins a chat episode by submitting a justification statement (Figure 3). Students can identify teammates by their nicknames and select any justification statement in the window to add their comments. A chat episode further develops when a player selects another player’s justification statement or comment and posts an additional comment (Figure 4). In this chat episode, the player can make any type of comment, either a statement addressing the original author’s justification statement, or comments to other players who have posted in this discourse episode. For clarification, the author is identified as the student who submits their decision and justification; the other students (teammates) respond to the author’s justification and/or additional comments within that thread. Since all the players in a game can see each other’s justification statements and comments, they are free to select and continue commenting in all the different chat episodes that were created at the end of the game. Players can continue commenting until the teacher or students end the game session.
Players also have the opportunity to add or remove game points from other players by providing a thumbs-up (adding one point) or a thumbs-down (removing one point) to any author’s justification statement or to any comment made by another player in any chat episode that is part of this game. The directions provided to the teachers, as well as several introductory videos, encourage students to add points to teammates who make quality comments and remove points from those who do not address the claim, evidence, or content of the article, or who make negative or non-constructive comments. This is referred to as peer-mediated chat and is the game-based strategy used to provide students with the opportunity to regulate the quality of the conversation or discourse in the game (Ault et. al., 2014).

Data Collection

The materials used in this study included the log files capturing information related to student game play of the Reason Racer game. Students accessed the game through the Internet and individual student performance data were recorded to a server database. The log files contain information about the scenarios that were utilized, decisions that were made about the claim (students determining if they should accept, reject, or withhold judgment about the claim), and student discourse about the scenario.

Data Analysis

This research was designed to explore differences in student discourse of scientific and socioscientific topics. During the fourth component of the Reason Racer game, students engage in discourse with peers about decisions regarding a claim and the content of the brief article. The Reason Racer Discourse Analysis Scoring Guide (RR-DASG) was created and refined in the current study by an iterative process involving multiple comparisons and discussions within the research team until a degree of consistency was reached. Refinements included collapsing item types and improving item descriptors for clarity, ensuring that they represented a coherent summary of author statements and interaction types. Once the RR-DASG could be applied with 85% reliability between the scorers, the research team began to review and score the discourse episodes.

Table 1. Summary of interaction codes for overall discourse episode

<table>
<thead>
<tr>
<th>Type of Interaction</th>
<th>Item Code within Interaction Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Substantive Interaction</td>
<td>1. Nonsense text, playful use of typing, no content</td>
</tr>
<tr>
<td></td>
<td>2. No interactions with another player, same player comments to self</td>
</tr>
<tr>
<td>Social Interaction</td>
<td>3. Social discussion, unrelated to content or game</td>
</tr>
<tr>
<td></td>
<td>4. General positive or supportive comments</td>
</tr>
<tr>
<td></td>
<td>5. General negative comments (mean-spirited)</td>
</tr>
<tr>
<td>Non-Specific Discussion</td>
<td>6. Agreement about or referring to content or process in a positive or agreeing way, agreeing with other teammates in the conversation</td>
</tr>
<tr>
<td></td>
<td>7. Disagreement about or referring to content or process in a negative or disagreeing way, disagreeing with other teammates in the conversation</td>
</tr>
<tr>
<td>Discussion Based on Components of Argumentation</td>
<td>8. Agreement with use of argumentation vocabulary or application evidence and reasoning</td>
</tr>
<tr>
<td></td>
<td>9. Disagree with use of argumentation vocabulary or application of evidence and reasoning</td>
</tr>
<tr>
<td></td>
<td>10. Questioning or asking for more information or explanation</td>
</tr>
<tr>
<td></td>
<td>11. Exploring different views with two or more viewpoints expressed, discussion between two or more players about claim, content, or process</td>
</tr>
</tbody>
</table>
The scoring process included the research team (a) reviewing the author statement and discourse episode in its entirety, (b) scoring the interaction of the overall discourse episode, (c) scoring the author’s decision, justification type and argumentation vocabulary use frequency, (d) scoring the teammate(s) type of comment and argumentation vocabulary use frequency, and (e) counting the frequency of likes/dislikes for the author and teammate comments as well as the number of teammates involved in the discourse episode, number of author comments and number of teammate comments.

The overall discourse episode was scored based on the type of interaction between students, such as no substantive interaction, social interaction, non-specific discussion, or discussion based on components of argumentation (Table 1). The author’s decision and justification, as well as the teammate(s) comment, were scored based on the type of statement provided. Table 2 identifies the statement categories as well as descriptors for types of responses in each of these categories.

<table>
<thead>
<tr>
<th>Type of Statement</th>
<th>Item Code within Type of Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agree or Disagree</td>
<td>1. Basic or simplistic, no explanation or description of why agree or disagree</td>
</tr>
<tr>
<td></td>
<td>2. Based on evidence in the scenario</td>
</tr>
<tr>
<td></td>
<td>3. Based on reasoning from the scenario</td>
</tr>
<tr>
<td>Questioning</td>
<td>4. General question about the claim or indicating that there was not enough information</td>
</tr>
<tr>
<td></td>
<td>5. Question about another’s statement based on the use of evidence</td>
</tr>
<tr>
<td></td>
<td>6. Question about another’s statement based on reasoning</td>
</tr>
<tr>
<td>Disagree with a Challenge</td>
<td>7. Disagreement and providing a new question, counterargument, or rebuttal,</td>
</tr>
<tr>
<td></td>
<td>8. Weighing both sides of the argument with no resolution</td>
</tr>
<tr>
<td></td>
<td>9. Indecisive based on evidence and williing to accept conflicting views, or</td>
</tr>
<tr>
<td></td>
<td>10. Withholding judgment based on conclusions about limited evidence, reasoning, or claim</td>
</tr>
<tr>
<td>Other</td>
<td>11. Unrelated to the content of the article or claim, social in nature,</td>
</tr>
<tr>
<td></td>
<td>12. Positive or affirmative and related to the topic but not specifically addressing the claim</td>
</tr>
<tr>
<td></td>
<td>13. Negative and related to the topic but not addressing the claim,</td>
</tr>
<tr>
<td></td>
<td>14. Assist other players on how to play the game or perform better</td>
</tr>
<tr>
<td></td>
<td>15. Correcting own or other’s grammar, spelling, word choice, etc.</td>
</tr>
</tbody>
</table>

The research team also identified each scenario as scientific or socioscientific, based on a mutually agreed upon definition of socioscientific and an understanding of the scenario content. Overall, 21 scenarios were accessed during the game play sessions, with five featuring socioscientific issues and 16 of scientific content. The scenario type and discourse episode scoring were used in this analysis exploring student discourse of scientific and socioscientific topics.

Descriptive statistics were used to determine the frequency of scenarios played, student decisions, and use of scientific argumentation vocabulary. The analysis included independent-samples t tests, using discourse episodes from all scenarios played using the interaction type and author statement or teammate comment type as the dependent variable with the factor being the type of scenario played; scientific or socioscientific.

### Results and Discussion

Data were gathered in 5,897 game play sessions from the six classrooms over a two-month period. This analysis addressed the chat episodes of the game for each student during his or her last game play day, totaling 937 discourse episodes. As student groups competed against each other in one game session, each student was the leader on a discourse episode, with any or all the other teammates participating in the chat. As a result, there might be four to six discourse episodes occurring simultaneously at the end of each game play session, with all teammates who played the game participating in one or more chat episodes. Students were not required to
engage in a chat but were generally encouraged by their teachers to comment on other student’s rationale statements and to use the thumbs-up or thumbs-down to reward teammates’ discourse.

The sample consisted of 937 discourse episodes. Approximately 74% of the episodes involved scientific scenarios (n=691), while 26% were socioscientific scenarios (n=246). Table 3 reports the frequencies and percentages associated with the scenarios used during all game play. The most frequently accessed scenario was a socioscientific issue, Energy Drinks? Don’t Waste Your Energy, and the least accessed scenario was Return of the Mammoth, also a socioscientific issue.

Table 3. Frequencies and percentages of scenarios accessed

<table>
<thead>
<tr>
<th>Type</th>
<th>Scenario Name</th>
<th>Scenario Topic</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scientific</td>
<td>1908 Russian Explosion</td>
<td>Meteoroids and comets</td>
<td>3</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Beam Me Up! (Keep It Clean!)</td>
<td>Teleportation</td>
<td>8</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>Carbon Dioxide Sponge</td>
<td>Absorbing carbon dioxide</td>
<td>44</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>Deep Oceans and Global Warming</td>
<td>Global warming</td>
<td>45</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td>Dogs Can Read Human Faces</td>
<td>Dog intelligence</td>
<td>89</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td>Elevator to Outer Space</td>
<td>Large-scale engineering projects</td>
<td>176</td>
<td>18.8</td>
</tr>
<tr>
<td></td>
<td>Graphene Valley</td>
<td>Can graphene replace silicon?</td>
<td>10</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>Leapin’ Lizards</td>
<td>Search and rescue robots</td>
<td>69</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>Panda Poop to the Rescue</td>
<td>New technology for biofuel production</td>
<td>8</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>That Shrimp Packs a Punch!</td>
<td>Super-strong materials from shrimp</td>
<td>12</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>The Artificial Leaf</td>
<td>A step toward energy independence</td>
<td>28</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>The Earth’s Two Moons</td>
<td>New theory explains features of Earth’s moon</td>
<td>51</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>The New North</td>
<td>Reversal of the Earth’s magnetic poles</td>
<td>57</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td>Was Einstein Wrong?</td>
<td>Challenging the speed of light</td>
<td>43</td>
<td>4.6</td>
</tr>
<tr>
<td></td>
<td>Weather Is One Big Headache</td>
<td>Relationship between weather and migraine headaches</td>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>Worm Glue? Give Me a Break!</td>
<td>Biomimicry leads to possible new bone glue</td>
<td>46</td>
<td>4.9</td>
</tr>
<tr>
<td>Socio-scientific</td>
<td>Are Fatty Foods Addictive?</td>
<td>Fat triggered endocannabinoids and overeating</td>
<td>27</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>Energy Drinks? Don’t Waste Your Energy</td>
<td>Risks associated with energy drinks</td>
<td>192</td>
<td>20.5</td>
</tr>
<tr>
<td></td>
<td>Mindless Eating</td>
<td>Nutrition</td>
<td>19</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Return of the Mammoth</td>
<td>Scientists trying to clone a mammoth</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>Violent Video Games and the Brain</td>
<td>Violent video game and aggressive behavior</td>
<td>7</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Author Justification Statements

After deciding to accept, reject or withhold judgment about a claim, the author provided justification for their decision. These justification statements were analyzed using independent-samples t test to evaluate the relationship between the scenario type and the author’s rationale. The independent variable, the scenario type, included two different levels: scientific and socioscientific. The dependent variable was the type of author
statement, as scored by the research team using the Reason Racer Discourse Analysis Scoring Guide. An independent-samples t test was conducted to investigate the types of argumentation discourse displayed by students when they engaged in chat as part of an online multiplayer game about both socioscientific and scientific topics. Results can be found in Table 4 and in the following paragraphs.

Table 4. Type of author justification statement

<table>
<thead>
<tr>
<th></th>
<th>Scientific Topic</th>
<th>Socioscientific Topic</th>
<th>95% CI for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agreement</td>
<td>M</td>
<td>SD</td>
<td>n</td>
</tr>
<tr>
<td>Agree or Disagree</td>
<td>.57</td>
<td>.49</td>
<td>396</td>
</tr>
<tr>
<td>Questioning</td>
<td>.00</td>
<td>.05</td>
<td>2</td>
</tr>
<tr>
<td>Disagree with a Challenge</td>
<td>.22</td>
<td>.41</td>
<td>152</td>
</tr>
<tr>
<td>Other</td>
<td>.20</td>
<td>.40</td>
<td>141</td>
</tr>
</tbody>
</table>

* p < .05

The test was significant, t(484) = -5.08, p = .00 for author justifications scored as “Agree or Disagree". Students playing the Reason Racer game using a socioscientific topic (M = .74, SD = .44), on average were scored as a general agree or disagree on their decision justification more often than those playing Reason Racer using a scientific topic (M = .57, SD = .49). The 95% confidence interval for the difference in means ranged from -0.24 to -0.10. The strength of the scenario type and student interaction scored as having a small effect size as assessed by η², accounted for 2% of the variance of the dependent variable. This category was scored for author justifications that were based on basic or simplistic explanations which may have included evidence and/or reasoning in the scenario.

The test was not significant, t(376) = -0.26, p = .80 for author justifications scored as “Questioning". Students playing the Reason Racer game using a socioscientific topic (M = .00, SD = .06), on average were scored as questioning on their decision justification equal to than those playing Reason Racer using a scientific topic (M = .00, SD = .05). This category was scored for author justifications that were based on general questions about the claim, or where students indicated that not enough information was provided, or that they had questions based on the use of evidence or reasoning.

The test was significant, t(519) = 3.19, p = .00 for author justifications scored as “Disagree with a Challenge". Students playing the Reason Racer game using a socioscientific topic (M = .13, SD = .34), on average were scored as disagree with a challenge on their decision justification less often than those playing Reason Racer using a scientific topic (M = .22, SD = .41). The strength of the scenario type and student interaction scored as having a small effect size as assessed by η², accounted for 1% of the variance of the dependent variable. This category was scored for author justifications that disagreed with the claim; provided a new question, counterargument, or rebuttal; indecisive, or withholding judgment about the claim.

The test was significant, t(534) = 3.35, p = .00 for author justifications scored as “Other". Students playing the Reason Racer game using a socioscientific topic (M = .12, SD = .32), on average were scored as other on their decision justification less often than those playing Reason Racer using a scientific topic (M = .20, SD = .40). The strength of the scenario type and student interaction scored as having a small effect size as assessed by η² accounted for 1% of the variance of the dependent variable. This category was scored for author justifications that were unrelated to the content, affirmative relating to the topic but not necessarily the claim, or negative related to the topic but not addressing the claim.

Overall, student authors within socioscientific topics supported their decision to accept the claim based on general agreement using basic or simplistic explanations, which may have included evidence, and/or reasoning from the scenario more often than student authors within scientific topics. Socioscientific topics also yielded fewer disagreements, student challenges, and unrelated justifications that did not address the claim.

Teammate Comments

The author’s decision to accept, reject or withhold judgment about a claim, as well as their justification for this decision fueled teammate comments within the discourse episode. These teammate comments were analyzed using an independent-samples t test to evaluate the relationship between the scenario type and the type of comment provided by the teammate(s). The independent variable, the scenario type, included two different
levels: scientific and socioscientific. The dependent variable was the type of teammate comment, as scored by the research team using the Reason Racer Discourse Analysis Scoring Guide. Results can be found in Table 5 and discussed in the following paragraphs.

Table 5. Type of teammate(s) comment in response to author justification

<table>
<thead>
<tr>
<th></th>
<th>Scientific Topic</th>
<th>Socioscientific Topic</th>
<th>95% CI for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>n</td>
</tr>
<tr>
<td>Agree or Disagree</td>
<td>.40</td>
<td>.91</td>
<td>273</td>
</tr>
<tr>
<td>Questioning</td>
<td>.21</td>
<td>.71</td>
<td>145</td>
</tr>
<tr>
<td>Disagree with a Challenge</td>
<td>.03</td>
<td>.20</td>
<td>23</td>
</tr>
<tr>
<td>Other</td>
<td>1.21</td>
<td>2.12</td>
<td>836</td>
</tr>
</tbody>
</table>

* p < .05

The test was significant, t(665) = 3.37, p = .00 for teammate comments scored as “Agree or Disagree”. Students playing the Reason Racer game using a socioscientific topic (M = .22, SD = .59), on average were scored as a general agree or disagree less often than those playing Reason Racer using a scientific topic (M = .40, SD = .91). The strength of the scenario type and student interaction scored as having a small effect size as assessed by $\eta^2$, accounted for 3% of the variance of the dependent variable. This category was scored for teammate comments that were based on basic or simplistic explanations which may have included evidence and/or reasoning in the scenario.

The test was not significant, t(546) = .33, p = .74 for teammate comments scored as “Questioning”. Students playing the Reason Racer game using a socioscientific topic (M = .20, SD = .56), on average were scored as questioning less often than those playing Reason Racer using a scientific topic (M = .21, SD = .71). This category was scored for teammate comments that were based on general questions about the claim, students indicating that there was not enough information, or questions based on the use of evidence or reasoning. The test was not significant, t(690) = 1.53, p = .13 for teammate comments scored as “Disagree with a Challenge”. Students playing the Reason Racer game using a socioscientific topic (M = .02, SD = .13), on average were scored as disagree with a challenge less often than those playing Reason Racer using a scientific topic (M = .03, SD = .20). This category was scored for teammate comments that disagreed with the author; provided a new question, counterargument, or rebuttal; indecisive, or withholding judgment about the claim. The teammates may have also been weighing both sides of the argument with no resolution or were indecisive based on evidence and willing to accept conflicting views.

The test was significant, t(320) = -2.72, p = .01 for teammate comments scored as “Other”. Students playing the Reason Racer game using a socioscientific topic (M = 1.82, SD = 3.29), on average were scored as other more often than those playing Reason Racer using a scientific topic (M = 1.21, SD = 2.12). The strength of the scenario type and student interaction, as assessed by $\eta^2$, accounted for 1% of the variance of the dependent variable. This category was scored for teammate comments that were unrelated to the content, affirmative relating to the topic but not necessarily the claim, or negative related to the topic but not addressing the claim. Teammates were also scored as “other” when they were assisting other players on how to play the game or perform better or correcting their own or another player’s grammar, spelling or word choice. Overall, socioscientific topics lead to student comments that were positive, helpful for other players, and supportive of their teammates more often than students discussing scientific topics. There was less general agreement/disagreement, questioning and challenging other players when students played Reason Racer with a socioscientific topic than when they played using a scientific topic.

**Discourse Episodes**

The discourse episodes occurring during the fourth part of the game (Figure 4) were analyzed based on the interaction type scored by the research team using the Reason Racer Discourse Analysis Scoring Guide. An independent-samples t test was conducted to evaluate the relationship between the scenario type and the student interaction within the discourse episode for each type of interaction. The independent variable, the scenario type, included two different levels: scientific and socioscientific. The dependent variable was the type of student interaction during the overall discourse episode. Results can be found in Table 6 and in the following paragraphs.
The test was significant, t(436) = 2.03, p = .04 for student interactions scored as “No Substantive Interaction”. Students playing the Reason Racer game using a socioscientific topic (M = .41, SD = .49), on average were scored as “No Substantive Interaction” less often than those playing Reason Racer using a scientific topic (M = .49, SD = .50). The strength of the scenario type and student interaction scored as having no substantive interaction, as assessed by $\eta^2_p$, accounted for 1% of the variance of the dependent variable. No substantive interaction was scored for discourse episodes having nonsense text, playful use of typing, no content or no interactions with another player. Upon further analysis of the discourse episodes, students playing Reason Racer using scenarios with scientific topics were scored as having no interactions with another player more often (46.8%) than students using socioscientific topics (40.2%). In other words, students playing the game with socioscientific topics played the game more often with teammates than alone.

The test was significant, t(400) = -2.72, p = .01 for student interactions scored as “Social Interactions”. Students playing the Reason Racer game using a socioscientific topic (M = .35, SD = .48), on average were scored as “Social Interaction” more often than those playing Reason Racer using a scientific topic (M = .26, SD = .44). The strength of the scenario type and student interaction scored as being a social interaction, as assessed by $\eta^2_p$, accounted for 3% of the variance of the dependent variable. Discourse episodes were scored as social interactions when the student engagement was mostly social discussion unrelated to the content or game, general positive or supportive comments, or negative, mean-spirited comments. Upon further analysis, students playing the Reason Racer game using a socioscientific topic were scored as having more general positive/supportive comments (27.6%) over students playing the game using scientific topics (19.7%). In addition, students in socioscientific topics were never scored as being negative or mean-spirited in their overall discourse episode interaction.

The test was not significant, t(435) = .15, p = .88 for student interactions scored as “Non-Specific Discussion”. Students playing the Reason Racer game using a socioscientific topic (M = .14, SD = .35), on average were scored as “Non-Specific Discussion” less often than those playing Reason Racer using a scientific topic (M = .15, SD = .48). Non-specific discussion included general agreement or disagreement about or referring to the content or process with other teammates.

The test was not significant, t(461) = .75, p = .46 for student interactions scored as “Discussion Based on Components of Argumentation”. Students playing the Reason Racer game using a socioscientific topic (M = .09, SD = .29), on average were scored as engaged in “Discussion Based on Components of Argumentation” less often than those playing Reason Racer using a scientific topic (M = .11, SD = .31). Discourse episodes scored as discussion based on components of argumentation included agreement/disagreement with the use of argumentation vocabulary or application of evidence or reasoning, questioning or asking for more information or exploration, or students exploring different views with two or more viewpoints expressed having discussion between two or more players about claim, content or process. Upon further analysis, students playing the Reason Racer game using a socioscientific topic generally were agreeing, questioning, or exploring different views using components of argumentation. They were never scored in disagreement with their teammates on socioscientific topics.

Overall, socioscientific topics produced positive social discourse relating to the issue being discussed within a competitive, yet collaborative, environment. While their conversations may have been less focused on the components of argumentation, students discussed potentially controversial topics in a civil and affirmative manner, supporting their own beliefs as well as the beliefs of their teammates.
Use of Scientific Argumentation Vocabulary

Throughout the discourse episodes, students were encouraged to use vocabulary related to scientific argumentation. This vocabulary included basic words such as claim, qualifier, and evidence, as well as words to describe evidence, such as data, fact, opinion and theory. Vocabulary also included terms relating to reasoning, such as authority and logic, and the application of reasoning such as if-then statements. Challenges to the claim were identified as counterarguments and rebuttals, while new questions were indicated by students using the terms why or how within the discourse episode.

Table 7 reports the frequencies and percentages of scenarios in which students used scientific argumentation vocabulary during their discourse episode. Scientific argumentation vocabulary was used by the author of the discourse episode 43.9% of the time for socioscientific topics (n=108) and 39.1% of the time for scientific topic scenarios (n=270). Teammates within the discourse episode used one or more scientific argumentation terms in 13.4% of the socioscientific topics (n=33) and 14.2% of the scientific topic scenarios (n=98). Authors within discourse episodes frequently used words such as accept, because, and fact.

<table>
<thead>
<tr>
<th></th>
<th>Scientific Topic (N=691)</th>
<th>Socioscientific Topic (N=246)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author using 1 or more</td>
<td>n=270</td>
<td>n=108</td>
</tr>
<tr>
<td>vocabulary term</td>
<td>39.1%</td>
<td>43.9%</td>
</tr>
<tr>
<td>Teammate(s) using 1 or</td>
<td>n=98</td>
<td>n=33</td>
</tr>
<tr>
<td>more vocabulary term</td>
<td>14.2%</td>
<td>13.4%</td>
</tr>
</tbody>
</table>

Conclusion

Counter to the findings from Linn and her colleagues, as well as others, demonstrating the ability to engage students in discourse and argumentation in scaffolded and controlled web-based spaces, argumentation discourse in an online chat-like environment was successfully implemented without programmatic scaffolding through the Reason Racer game (Jeong & Joung, 2007; Linn, Clark, & Slotta, 2003; Linn & Eylon, 2011). The salient characteristics identified by Soller et al. (2005) as common to applications that support collaboration include a shared work-space that supported a social awareness of teammates, a chat function allowing for open-ended interactions, delineated roles, problem-solving actions, and graphical visualizations of performance. These were present within the game environment. The graphical displays of the interactions were similar to social sites such as Facebook, Twitter, or other chat environments. They did not provide students with an extensive metacognitive perspective of their actions; instead, allowing for productive discourse to take place.

Soller et al. (2005) also suggested that a key component of the online system was the manager who provided feedback, remedial actions, or helped students with their online behaviors. The game, particularly the chat environment, also provided feedback, but from peers, rather than a manager. While the students were aware that the teacher could also see their comments, it could be argued that the peer mediation also had a mitigating nuance on discourse.

Scoring students’ discourse involves components of argumentation, as outlined by Toulmin, as well as other components, such as an alternative strategy used to quantify what occurs during the process of scientific argumentation focused on discourse characteristics as either sense-making or persuasion (Berland & McNeill, 2010; Berland & Reiser, 2011). Through analyzing student discourse episodes, argumentation components, as well as student understanding of argumentation, can be identified. While some results were significant for the analysis categories, effect sizes remained small for these interactions. In general, students engaged in discourse of socioscientific topics had positive, supportive, and civil dialogue with their teammates.

Limitations of the Study

We believe there are two possible limitations to this study. First, there were a limited number of scenarios accessed for student game play. The Reason Racer game contains 40 scenarios containing scientific and socioscientific topics. The data set analyzed only included 21 of these scenarios, or a little over half of the scenarios. Most the 21 scenarios represented scientific topics (n=16), while 5 scenarios were about socioscientific topics. While this imbalance in scenarios can be addressed through statistical methods, it is clear that more scenarios were accessed for scientific topics rather than socioscientific topics.
Second, because the research team was only accessing the game play log files and did not have interaction with the classroom teacher regarding the selection of scenarios, it cannot be determined if the scenarios were assigned to the students as part of a classroom lesson, or if students were able to choose the scenario they played. Student choice could make a difference in discourse engagement for the different scenario types. Similarly, students played at least one game, but up to five game sessions that were included in this analysis. Therefore, if students had more opportunity to play and discuss across multiple scenarios it might have an impact on discourse engagement.

**Recommendations**

The *Reason Racer* game can be used to begin the process of engaging students in discourse and argumentation. The chat feature within the *Reason Racer* game is effective to monitor discourse in scientific and socioscientific topics, providing feedback and models for students to support the development of the skill. Recognizing that it may be difficult for middle school students to grasp what is “fun” about engaging in argumentation, this study investigated the differences in student-level discourse between socioscientific and scientific topics as discussed through an online game. This study finds that students can and will engage in productive and positive discourse through socioscientific topics. These results add to what is known about using online, educational games in the classroom for collaborative discourse.

**Acknowledgements**

This research was supported by the National Science Foundation under Grant #1019842. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

**References**


---

**Author Information**

**Jana Craig-Hare**  
University of Kansas Center for Research on Learning  
U.S.A.  
Contact e-mail: janach@ku.edu

**Marilyn Ault**  
University of Kansas Center for Research on Learning  
U.S.A.

**Amber Rowland**  
University of Kansas Center for Research on Learning  
U.S.A.