PREDICTION GAMES: ENCOURAGING ENGAGEMENT WITH DATA

A Dissertation

by

GABRIEL SERGE DZODOM

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Frank M. Shipman, III
Committee Members,	Richard Furuta
	James Caverlee
	Susan Pedersen
	Cathy Marshall
Head of Department,	Dilma Da Silva

December 2020

Major Subject: Computer Science

ABSTRACT

Prediction games are data-driven games modeled after fantasy sports. They are designed to motivate players to explore, analyze, and develop their own understanding of large data sets. A survey of fantasy sports players explored the data analysis and decision-making processes for prediction games, and identified news and social media content as a common source of information for players. Based on the survey results and experiences with an early prediction game prototype, a system's framework for prediction games was developed and a game in the climate domain was implemented: Fantasy Climate. Fantasy Climate asks players to select a location among a set of choices based on their expectations of how upcoming weather will vary from historical norms. User studies of Fantasy Climate revealed that the in-game presentation of domain-related news had a significant effect on player engagement and performance. Additionally, in-game asynchronous social interaction methods made the game more engaging and helped players with their data analysis and prediction making strategies.

Prediction games are meant to be valuable in formal and informal educational settings. But creating a prediction game normally requires significant software development. To enable non-programmers to author prediction games, we designed and developed the Activity Creation Wizard (ACW). The ACW includes a thorough help system, a template component for reusing prior game designs, and tools to automate repetitive and tedious tasks. An evaluation of the ACW showed no background knowledge was required to author a prediction activity. The help system met most of the

participants' information needs, templates were found useful by many, and automation reduced the time taken for repetitive tasks. Reasoning about the effects of their choices on gameplay was noted as the primary challenge during the authoring task by several participants. The evaluation identified alternative approaches to authoring that challenged the current design, including the potential value of co-dependent customizations and collaborative authoring. Finally, participants also were asked to create a prediction game in the domain of their choice. Interviews regarding these creations revealed that educational, social, and socio-cultural factors play an important role in what makes prediction games engaging.

DEDICATION

To my sister Carine Megang Nzodom, MD.

ACKNOWLEDGEMENTS

Above all, I would like to express my deepest gratitude and appreciation to Dr. Frank M. Shipman III for his support, guidance and patience during my graduate studies.

I would also like to thank my committee members Dr. Richard Furuta, Dr. James Caverlee, Dr. Susan Pedersen, and Cathy Marshall for their guidance and support in improving the quality of my research.

I also want to extend my gratitude to all the mentors in my life that participated in my professional and personal growth and, helped carve the path that led me to this point. I would especially like to recognize Dr. John Keyser, Dr. Greg Turk, Dr. Chao Lu, Dr. Alfred Weaver, Dr. Donald Gooden and, Martha Ralston.

Finally, I would like to thank my family and friends for their encouragement, patience and love.

CONTRIBUTORS AND FUNDING SOURCES

Contributions

This work was supervised by a dissertation committee consisting of the advisor: Dr. Frank M. Shipman (the advisor), Dr. Richard Furuta, Dr James Caverlee, and Cathy Marshall of the Department of Computer Science and Engineering and, Dr. Susan Pedersen of the Department of Educational Psychology.

Meghanath Reddy Junnutula of the Department of Computer Science and Engineering developed parts of the system described in section 5.3 and conducted the user study reported in 6.1. This work was published in his Masters thesis in 2015.

Akshay Kulkani of the Department of Computer Science and Engineering developed parts of the system described in section 5.2 and conducted the user study reported in 6.2. This work was published in his Masters thesis in 2016.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was partly supported by the Diversity Fellowship from Texas A&M University. This work was also made possible in part by the National Science Foundation under Grant Number 1816923. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Science Foundation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

1. INTRODUCTION[*]

Big Data is a term that usually refers to datasets that are large, complex (unstructured, varied), and dynamic (continuously expanding) [26,39]. The popularity of Big Data rides on advances in digital sensors, communications, computation, and storage that have enabled massive collections of data about ourselves and our environment [37]. This "Datafication", collecting data about most anything under the sun, is of great value to business, science, government, and society [38]. The website data.gov [61], a public repository for data from governmental agencies and institutions, supports open access to government data. Regardless of their sources (government, science, business, or people), the rich variety of available data provides new opportunities for people to become better informed citizens [3,62]. For example, the National Oceanic and Atmospheric Administration (NOAA) or, the National Aeronautics and Space Administration (NASA) collect, maintain and publish data regarding the Earth's current and prior climate. However, it is unreasonable to expect most individuals to seek out this climate data and analyze it for their understanding without external motivation.

Interactive visualization and simulation techniques facilitate data interpretation, increasing the accessibility of large data sets [19]. An increasing number of websites provide open datasets to produce informative visualizations or serve as a hub for data

[*] Part of this chapter is reprinted with permission from the paper: © 2016 ACM. Gabriel Dzodom and Frank Shipman. 2016. Data-driven Prediction Games. In proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA' 16). Association for Computing Machinery, New York, NY, USA, 1857 – 1864. DOI:https://doi.org/10.1145/2851581.2892546

visualization projects [36,63,64]. Informationisbeautiful.net maintains a gallery of interactive visual artifacts that focuses on the environment and climate (Figure 1) [65]. However, there are limitations and risks associated with supporting data interpretation via interactive visualizations and simulations. For one, the gulf between the user mental model and the designer mental model may lead to confusion, misleading interpretations, or information overload [5]. In other words, the user is at the mercy of the designer's interpretation and view of the data. Second, the visual representations or the simulation models are often built on fixed dataset(s). Hence they may not reflect local or recent events that can motivate user attention [7], which can limit learning [18]. We have been investigating an alternative approach to motivate people to engage with open datasets bringing together domain data, domain-related information resources, and tools to make sense of this information in a game context. We coin these games: prediction games.

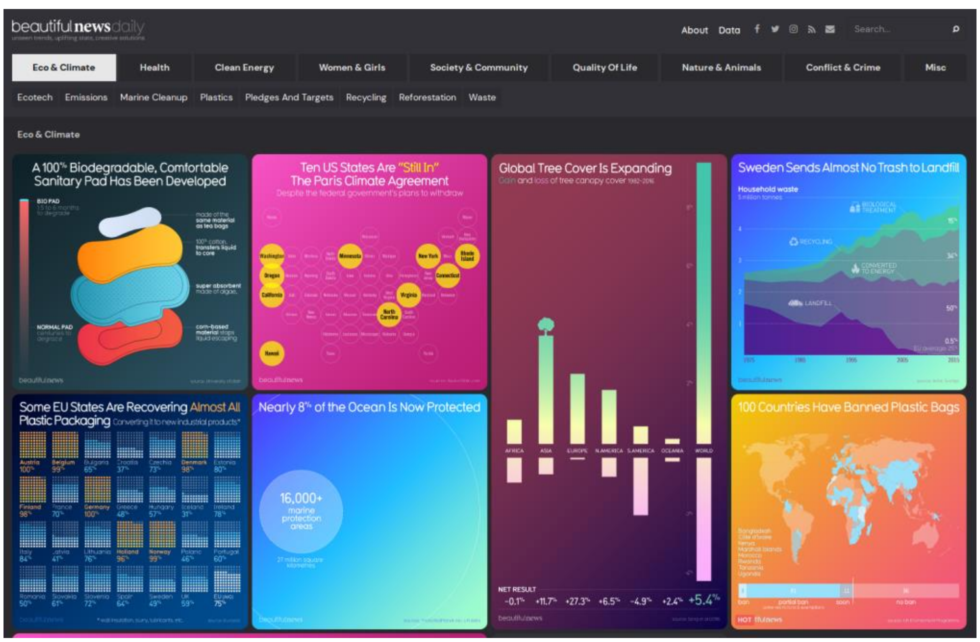

Figure 1. Visualizations on climate on informationisbeautiful.net. Reprinted from [65]

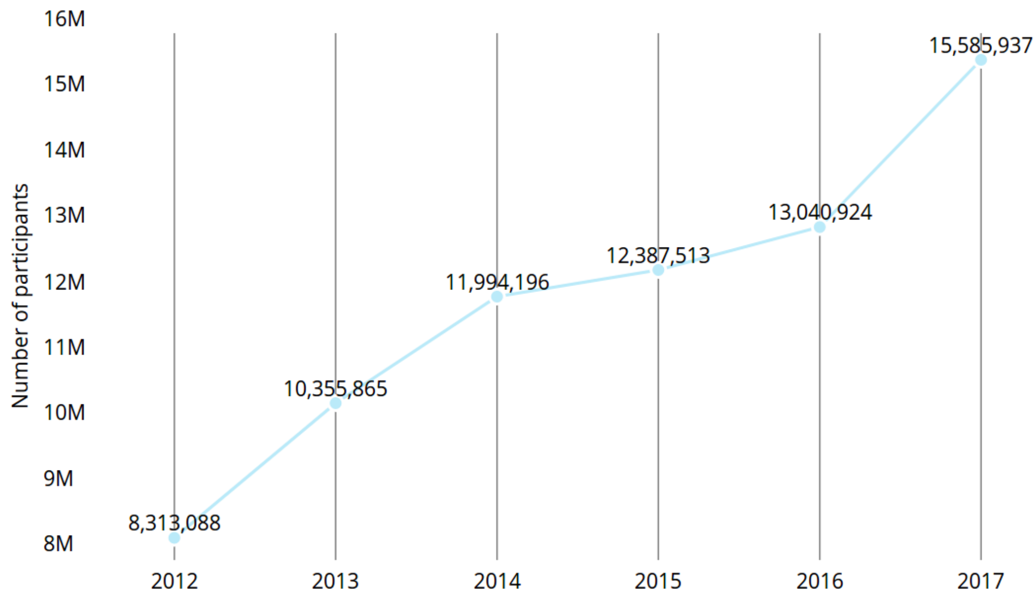

Figure 2. Fantasy sports participation over the years. Reprinted from [66]

Our exploration of prediction games is motivated by the success of fantasy sports. Fantasy sports have players act as if they are the manager of a team of athletes [53]. The player competes against other fantasy team managers directly or in a league. Similar to real team management, players decide which athletes to start and trade athletes with other fantasy players. Each player's score is based on statistics about the performance of their athletes, generated by a single athlete or a team during a real sporting event. Thus, up-to-date knowledge of the sport in question is valuable for these games. Fantasy sports are very popular. According to Nielsen Scarborough, fantasy sports participation in the U.S. increased from approximately 8.3 million in 2012 to 15.6 million in 2017 (Figure 2) [66]. Furthermore, studies show that fantasy sports motivate players to explore and interpret large data sets in order to generate better predictions

[11,12,30,48]. As a result, the players learn more about the sport. Therefore, fantasy sports serve as a success model for our design of prediction games.

Prediction games encourage players to examine historical data and domain-related information resources to make predictions about future events. As a consequence, prediction games can be developed for any domain with data archive about past activities and upcoming events to predict. We envision prediction games as being of value in formal and informal educational environments. It has been well established that games can be effective at engaging learners hence improving learning and understanding [25,28,45].

This research explores how to design prediction games and how to support their authoring. In particular it asks:

- what capabilities are necessary and valuable for prediction games, and
- how to support the creation and the customization of prediction games

activities with very little or no programming?

To answer these questions, we first design, develop, and evaluate a prediction games system which then informs the design and development of an environment for authoring prediction games activities.

This dissertation is organized into two major sections: (a) designing prediction games, and (b) designing an authoring environment for prediction games activities. The first section begins with a discussion of related work studying fantasy sports and building games in other domains that share aspects of the fantasy sports. This is followed by the results of a survey of fantasy sports players about their data gathering and analysis practices. An understanding of player practices is important to designing engaging prediction games that also meet the goal of increasing data analysis skills. After this is the description of an early prototype prediction game, Fantasy Forecaster, developed to gather system requirements and early user feedback. Results from the survey and experiences with Fantasy Forecaster lead into a conceptual framework of such systems and the design of Fantasy Climate, a prediction game in the climate domain. Next, we report on two evaluations of Fantasy Climate that focus on features that are known to affect engagement in fantasy sports but that are rarely included in prior educational games modeled after fantasy sports: (1) the use of alternative communication mechanisms and (2) the effect of News presentation on engagement and prediction accuracy.

Building on the conceptual framework of prediction games described in the first section, the second section explores the support of authoring prediction games. It begins with a discussion of prior work on domain-oriented design environments and instructional design environments. After this is a description of the Activity Creation Wizard (ACW): an authoring environment that guides an activity author (e.g. an instructor) through the creation and customization of a prediction game like Fantasy Climate. Finally, we report on an evaluation of the ACW for Fantasy Climate that examines how different components affect the authoring of prediction games activities, what aspects of prediction game authoring users find most challenging, and the variety of activities generated when participants are asked to design a new prediction game.

2. FANTASY SPORTS AND OTHER DATA-DRIVEN GAMES†

This research into prediction games builds on the results from analyses of fantasy sports practices and motivations, prior efforts using the fantasy-sports model to develop activities in non-sport domains, and the design of data-oriented games more broadly.

2.1. Fantasy Sports

An understanding of the motivations of fantasy sports players and the effect of features of the gaming environments and their communities on gameplay is valuable in designing prediction games. Fantasy sports are a significant component of the commercialization of sports. As a consequence, the goal of many prior studies of fantasy sports player activity has been to build a profile of the average user as a potential purchaser of sports related goods and services. This focus on the relationship between fantasy sports and sports consumption shows that fantasy sports players are more likely to attend live sports events and tend to consume mass media (television, newspaper, radio, etc…) more than other sports fans [9,10,32,42,46,56].

One aspect of prior studies relevant to prediction games is the understanding of player motivation. Lee et al. identified five key motives for fantasy sports involvement: hedonic experience, escape, prize, bonding with friends or family, and social interaction [35]. Ruihley and Hardin classified three stronger (fanship, competition, and social

† Part of this chapter is reprinted with permission from the paper: © 2016 ACM. Gabriel Dzodom and Frank Shipman. 2016. Data-driven Prediction Games. In proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA' 16). Association for Computing Machinery, New York, NY, USA, 1857 – 1864. DOI:https://doi.org/10.1145/2851581.2892546

sport) and three weaker (fan expression, ownership, and escape) motivations. They reported that competition, achievement, and surveillance motivations were positively correlated with overall satisfaction while competition and camaraderie motivations were positively correlated with future intentions to play fantasy sports [49]. Farquhar and Meeds [16] took another approach by building five profiles of fantasy sports users based on their motivations (casual players – with social and entertainment motivations; skilled players – entertainment, achievement, and social; isolationist players – arousal and social; trash talkers – arousal and social; and formative players – entertainment, achievement, and intellectual challenge). Taken together, these results point towards five common motivations: (1) engagement with sport, (2) escape and entertainment value, (3) prizes (which are usually financial), (4) social interactions and (5) intellectual challenge and competition.

Researchers have also explored the implications of fantasy sports demographics (gender and race) and interactions. Davis and Duncan studied fantasy league activity through a combination of observations, textual analysis, and focus group responses. They found that fantasy sports reinforce traditional gender and racial roles in sports [8]. Ruihley and Billings examined the consumption behaviors and motivation of men and women in fantasy sports. In five of their seven measures: arousal, entertainment, escape, self-esteem, and surveillance, their results reveal that men and women had similar motivations for participating in fantasy sports. They noted that the significant measured gender difference for the remaining two measures (enjoyment and passing time) might be due to the fact that men in their study consumed 10 hours more sports weekly than women, and that they ranked sports higher among leisure time activities than did women [48].

Ruihley and Hardin analyzed fantasy sports message boards to understand the effects of socializing on the players' experience. Their study found that social interaction through message boards is vital to and enhances the fantasy sports experience. Message board users were more satisfied and more likely to continue playing fantasy sports than non-message board users. Some of the principal motivations behind the utilization of the message boards were trash-talking, advice seeking, and general discussion of fantasy sports [50].

More related to the data analysis goals of our work, Smith et al. examined messages from forum discussions associated with fantasy basketball to understand the decision-making strategies of the users. Their findings indicated that the majority of fantasy basketball users opted for "satisficing" strategies rather than optimal ones. The reason, the authors explained, is that players have a limited amount of time to collect enough information for making these decisions [54]. In another study, Ruihley and Hardin found that in general, fantasy sports sites do not meet the informational needs of their users. They reported that the typical fantasy sports player employed additional web sites and television programming for gathering news, expert advice and athlete statistics [47]. Hirsh et al. confirm these results in their investigation of information-seeking behaviors and needs in fantasy football and fantasy NASCAR. In addition, they uncovered that although fantasy sports players valued predictive and trend statistics, there was a threshold where the volume of data hindered their experience. Moreover, they also found that social media is becoming an important information gathering tool of fantasy sports users [29]. Finally, Wohn et al. synthesized all these results in their work to understand the information-seeking and decision-making behaviors of daily fantasy sports (DFS) players. Their findings revealed that players of DFS thought that success partially depended on spending time collecting and assessing information through a variety of socially- and technologically-mediated research practices. In general, there was no single tool nor practice that suited the needs of players but a variety of practices across individuals [59].

These studies of motivation and communication in fantasy sports have informed our efforts. They indicate that there are several factors that affect motivation, some of which will be present in non-entertainment domains and others which will not. Also, they point out the importance of in-game data availability and social activity.

Missing from the earlier studies is an effort to determine the variety and quantity of data used by players to make game decisions and the tools and time used to interpret this data. As a consequence, we report on a survey of fantasy sports players exploring engagement with data in fantasy sports to better understand the game design attributes that make fantasy sports successful and their current limitations.

2.2. Fantasy Sports Game Model in Other Domains

The success of fantasy sports has sparked interest in applying its game model to other domains. FantasySCOTUS is a Supreme Court fantasy league for law students and high school students that encourages learning about constitutional law and current

Supreme Court decisions [4]. The prediction page for FantasySCOTUS is shown in Figure 3.

Figure 3. FantasyScotus prediction page. Reprinted from [67]

The game play involves predicting how each member of the Supreme Court will vote on a case. Points are awarded based on correct predictions. To do well in the game, the player should know the voting history of each member, the features of prior and current cases, and what politics are associated with the current case. FantasySCOTUS has had more than 10,000 users and has been featured in the New York Times and

CNN.com [67]. Prediction games have emerged in other domains including movie ticket sales [68,69], congressional decisions [70], and investments and property [71,72].

Related to climate, the Weather Channel held Fantasy Snowfall Leagues in 2013 and 2014 to spark interest in weather patterns and their prediction. In this league, players would form teams by selecting from a list of given cities. The previous year's average snowfall and average annual snowfall for every city are made available to the players. To play the game, the players predict cities with most snowfall in a descending order. At the end of each week, the actual snowfalling the cities is compared to the teams' predictions. The teams' points awarded depend on the accuracy of their predictions [14].

The above applications of the fantasy sports model have been developed with the intent to motivate learning but have not explored how alternative designs affect player engagement and activity. Our work explores the effects of alternate designs of prediction games. We also present a prediction game framework that can be adapted to any domain where there exists historical and real-time data. Finally, we introduce games that score results based on both historical data and real-time data make them effective at motivating learning about long-term trends in data.

2.3. Other Data-driven Games

Friberger et al. developed a taxonomy of data games that emphasizes how alternative data sources, data selection techniques, and data transformations enable different types of games [22]. They propose to use open and linked data to procedurally generate game content. For example, in their game Open Data Monopoly [23], the game board is generated based on player-defined formulas for defining prosperity based on real demographic and geographic data. However, the authors caution that game designers must be careful of transformations that “lead to an unacceptable loss of veracity in relation to the original source” [6,22,57]. When data is used to generate a simulation within a game, the resulting model may caricature and simplify the actual processes. In data-driven prediction games, the content examined by players is data from the real world and the challenge is to predict real future values – there is no abstraction or inexact modeling of the domain in question.

3. INFORMING PREDICTION GAMES DESIGN

Two formative activities were undertaken to better understand likely reactions to prediction games. First, we fielded a survey of fantasy sports players to better understand the data collection and analysis activities they undertook as part of the game. And second, we created a prediction game for weather forecasting to understand the system challenges to acquiring data, managing player accounts and leagues, and gathering initial reactions.

3.1. Fantasy Sports Players Data Practices[‡]

Fantasy sports is the success model for prediction games but prior studies of fantasy sports have not explored the data gathering and analysis practices of players. The decisions made by fantasy sports players may vary based on the sport and type of fantasy league. The first decisions are often during a drafting phase at the start of a league that occurs before the start of the sporting season. Drafts for fantasy sports leagues are part of leagues where each athlete may be part of only one team's roster. As such, players have to prioritize their choices based on their personal expectations of performance during the season and their understanding of the athlete's desirability to other players (e.g. are they an obvious star or a sleeper.) Once initial rosters are set, players make decisions about which of their athletes to start each game period (e.g. based on the athlete's opponent,

[‡] Part of this section is reprinted with permission from the paper: © 2014 ACM. Gabriel S. Dzodom and Frank M. Shipman. 2014. Data-driven web entertainment: the data collection and analysis practices of fantasy sports players. In proceedings of the 2014 ACM conference on Web science (WebSci' 14). Association for Computing Machinery, New York, NY, USA, 293 – 294. DOI:https://doi.org/10.1145/2615569.2615649

injuries) and whether to trade an athlete on their roster to another player or to drop them for an unclaimed athlete. Data about prior performance is known to play a significant role for all these decisions.

Given the goal of improving data skills and domain knowledge, understanding what and how much data players collect, how the data is used, and how much time is spent on game activities will provide insight for the design of prediction games. We conducted an Amazon Mechanical Turk survey of the practices of self-identified fantasy sports players to better understand the players' prediction-making practices. The survey included 7 multiple choice demographic questions, as well as 13 multiple-choice questions and 11 open-ended questions about fantasy sports experiences and practices. Appendix A describes the survey questionnaire. The following research questions drove the development of the survey:

Q1: What general strategies do players use to select athletes for their teams?

Q2: What are players' data collection and analysis practices?

(a) Where does the data and information used come from?

(b) What tools are used for data analysis?

Q3: How do players scope the data they include in their decision process?

Q4: What is the time commitment and activities of players?

Q5: What aspects of fantasy sports interfaces and activities they find difficult?

The results that follow are based on 160 survey responses. Most participants reported engaging in significant data collection and analysis as part of their athlete selection process. The majority (61%) of the responses mentioned data analysis as part of their process for selecting athletes. This was illustrated in the following statements: "*I buy a magazine and check the top rated players at each position. I also upload the previous seasons' stats into an Access database and play around with these figures trying to identify trends that are not normally in a 5x5 league that will help my draft choices*", and "*Stats is the way to go, people that go with their gut end up getting their feelings hurt.*" Players that did not report data analysis often mentioned reducing their effort by using their game's autodraft feature (where the game picks the "best" available athlete for the player based on simulations predicting athlete performance) while a small number mentioned avoiding online simulations in order to not turn the game into a statistics exercise. For example, one participant elaborated: "*I never simulated a thing, I just played for fun, pretending to be a coach from the old days when I was growing up. Trying to rely on instinct and not simulations and data analysis like happens now in sports. For me that stuff ruins the fun of the game.*" These latter players still reported using television and online stories to inform their decisions.

Where does this data and information come from and how much was used? When asked about where the information used during player selection came from, respondents more often preferred online information resources that were external to the fantasy sports game (social networks, league and sports news sites, etc.) over the ones that were internal to the game (news feed, message boards, etc.) Especially, news sites were primary sources of information for about half of them. Furthermore, many respondents reported using statistics about past athlete performance or predictions of future performance from software or experts. When asked about the number of different data fields (i.e. individual statistics about athletes) used during athlete selection, the majority of respondents considered seven or fewer as shown in Figure 4. When asked how they made sense of this data they collected, Figure 5 shows the most agreed upon answer was a combination of tools built into the particular game site and external tools (e.g. MS Excel, sheet of paper.)

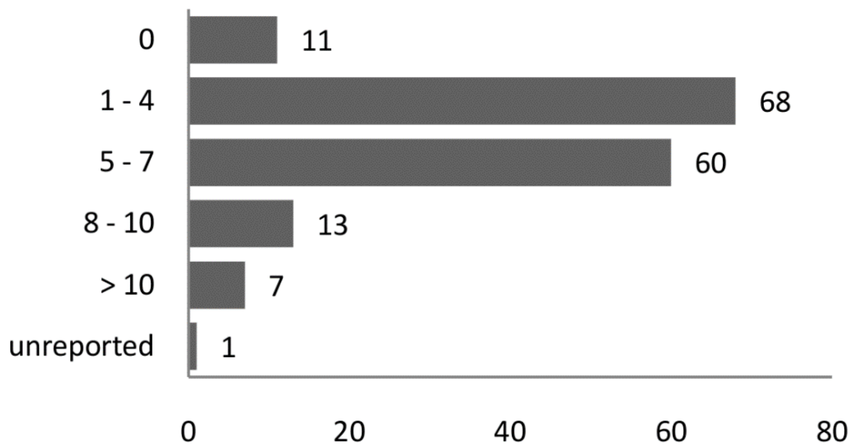

Figure 4. The number of data fields employed during the athlete selection process. Reprinted from [13]

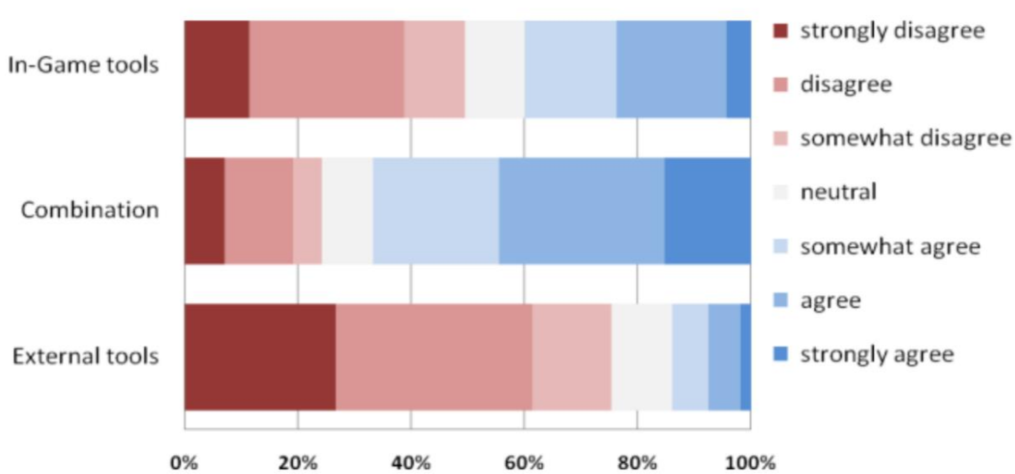

Figure 5. Location of tools used for data analysis.

One player explained use of external data analysis thus: "*The tools supported by the fantasy game were only geared toward an audience of a specific skill level. In order to step up my game, I needed to broaden my resources*." The use of external tools does not mean these players were all using spreadsheets and statistical analysis packages (although some were): "*I used a combination of both. The game-supported tools were used to figure out which players not only I had drafted, but others players had drafted as well. In addition, the game-supported tools allowed me to figure out which positions I had filled and which positions I had not. I used an external tool, which in this case, was a sheet of paper. I printed out my Big Board and ranked the players in who I thought was best. As a player would get selected, I would cross out his name on that sheet.*"

Once their team is initially set for a scoring period, almost all the respondents reported returning to the fantasy sports site at least occasionally to get updates on athletes, to monitor their team's score and its status in the league, and to interact socially with other players. In general, most of them visited the game site between two to seven times per week (Figure 6) where during each visit they spent between 15 minutes to an hour (Figure 7.)

While the primary purpose of this study was to explore data collection and analysis practices, questions also elicited motivations for playing. While social motivations were identified alongside others in prior studies of player motivation, many responses in this study indicate that this is the most important reason for a considerable number of players. For example, maintaining bonds with family and friends, "*it is a way* *to connect with my father who loves sports and also plays*", or workplace relations, "*To take part in an office activity with co-workers and not look like someone who doesn't want to participate in social activities in the workplace.*"

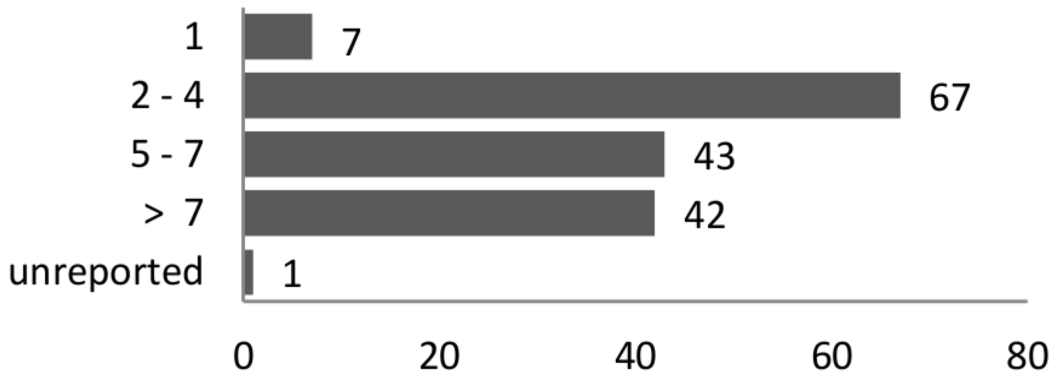

Figure 6. Number of game sessions per week. Reprinted from [13]

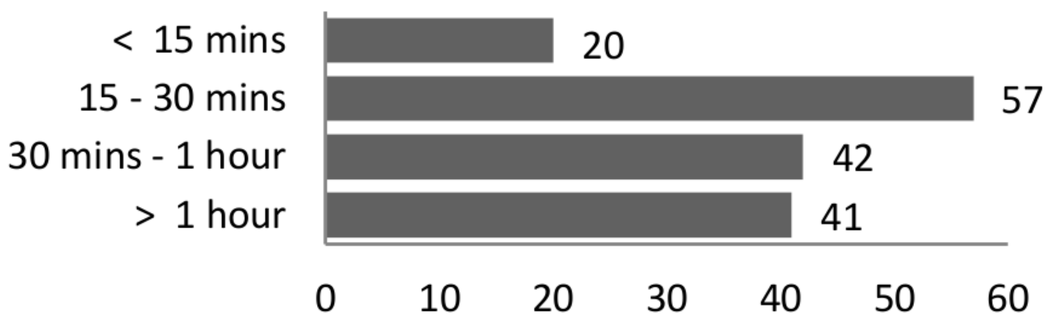

Figure 7. Length of average session. Reprinted from [13]

When asked about their frustrations and issues with fantasy sports, one of the main limitations reported was the lack of control over the data presentation and organization during analysis. For example, the inability to sort or filter data: "*I like to sort differently than some websites. I'm limited by how the sports leagues decide to rank players, so I prefer to rank them and look at different time frames and upcoming* *opponents.*" Finally, some participants wished for better support. Particular examples were tools enabling data comparison/visualization, and data export capabilities.

These results informed our design of prediction games by providing a sense of how much data most fantasy sports players work with and how they analyze this data. Participants also identified going outside the game to gather additional forms of data and to generate personalized analyses and visualizations. Thus, prediction games need to facilitate data collection by integrating information resources and support good data visualizations that enhance understanding and pattern identification. They should also relinquish control over data by giving players the opportunity for export data to an open format or providing a way for the players to navigate to the data source. Finally, the game needs to provide a social environment where players can build and improve relationships and seek help from their peers.

3.2. Prototype: Fantasy Forecaster[§]

The next step in designing prediction games was to identify fundamental system requirements and to gather initial user feedback. Hence we developed and evaluated a prototype: Fantasy Forecaster. The objective of the game is to make forecasts for U.S. locations based on historical weather data about these locations. Each player predicts the daily low/mean/high temperature, low/mean/high humidity, precipitation, cloud cover and events. They are scored based on the accuracy of their predictions when compared to the actual observed values from Weather Underground (www.wunderground.com.)

[§] Part of this section is reprinted from Weather Data Gamification, by Rohit Gargate, Texas A&M master's thesis. http://hdl.handle.net/1969.1/151174. Copyright 2013 Rohit Gargate.

Fantasy Forecaster supports players' prediction activities through visualizations, summarizing statistics, and information resources (i.e. news articles or facts) from various sources (e.g. NOAA, Weather Channel). As with fantasy sports, Fantasy Forecaster features leagues (i.e. groups of players that compete and see one another's performance) and teams except that in this case, a team consisted of U.S. cities instead of athletes.

To play the game, players first build a team by selecting a set of U.S. cities for which they will make predictions. Figure 8 shows the prediction submission interface. Below the top menu bar is a map where all the cities in the player's team are marked. The map panel footer displays the current submission date and the prediction date. On the left half of the page below the map panel, the cities are listed by tabs (one city per tab). For each city in her team, the player would enter a predicted value for each weather property (e.g. precipitation). The right side of the interface presents past data and statistics for the city in focus. Below the statistics panel is a visualization of quantitative weather data properties over 10 years to aid the players in identifying patterns. Once the observed real-world data is available and the predictions are scored, players can navigate to the performance page to see a report of how close their predictions were to the observed values. Finally, players can view their score and rank on the leaderboard.

A formative user study of Fantasy Forecaster was performed to identify issues in the software environment and gameplay and to identify which features were viewed as most helpful. Ten participants were recruited from the university community, mostly graduate students across a broad range of disciplines. The participants were asked to build a team (select a set of cities) and to submit predictions each day for seven days. After the week of game play was over, a post-game survey gathered feedback to open ended questions about participants' experience.

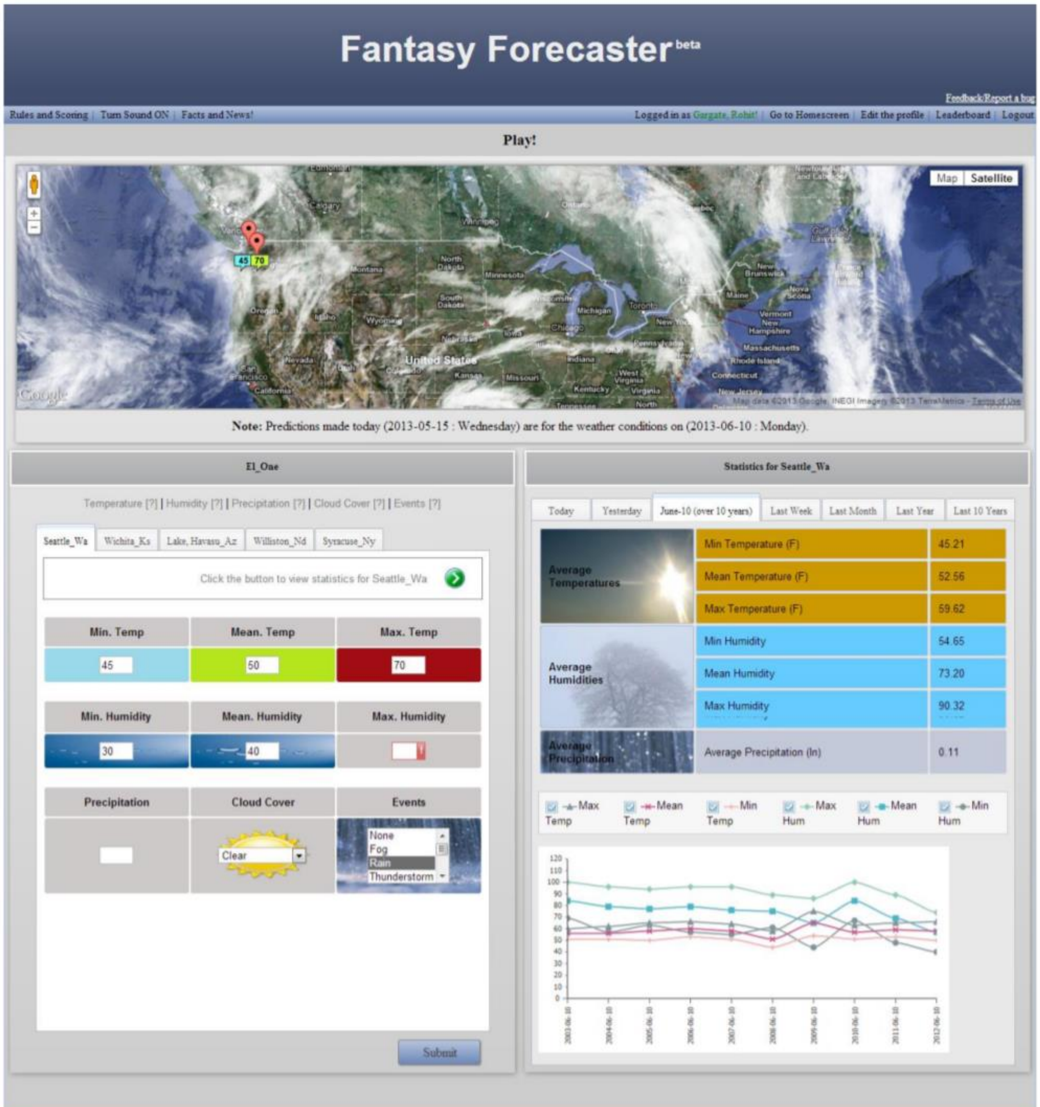

Figure 8. Fantasy Forecaster interface. Reprinted from [24]

Results from the survey showed that Fantasy Forecaster expanded the participants' curiosity about weather patterns around the world. When asked whether they were more engaged than usual with weather data and news during gameplay: "*Absolutely! Exponentially more!*" Their opinions of the news and facts page were also generally positive: "*I liked the 'Did you know' feature. I did not know any of those!*" or "*I like them, very interesting. The actual facts were fun.*" As a result, the participants reported learning new concepts and facts: "*...Learned about cloud cover units and how they are measured. Also thought Death Valley had the least rainfall in the world. However I learned that it is actually Atacama Desert*" and "*I thought that Seattle received more rain compared to New York but it is actually the reverse.*" Finally, the players' favorite features were the visualizations (e.g. graphs) of the weather data and the tabular representation of their performance statistics.

The development of Fantasy Forecaster identified three relatively-independent components of a prediction game engine: a content collection and management system to gather and maintain data and information, elements to support data exploration and analysis (e.g. visualization tools), and elements to make, edit, score and present the results of predictions.

Overall, the development and formative study of Fantasy Forecaster provided valuable insights into an appropriate organization of a predication game engine, necessary and valuable capabilities of such an engine, and initial indications as to the potential of prediction games outside of entertainment domains. The prediction game framework described in the next section and the Fantasy Climate prediction game that follows build on these results.

4. PREDICTION GAMES FRAMEWORK

Three software components, the data collection and storage system, data exploration tools and prediction facilities, were identified as crucial during the development of Fantasy Forecaster. Additionally, the survey of fantasy sports players identified the value of information beyond the quantitative data used for scoring the predictions (e.g. news) and the importance of social interactions. Our prediction game framework is structured around the types of content necessary for these capabilities, the sources of this content, and the interfaces players use when interacting with that content.

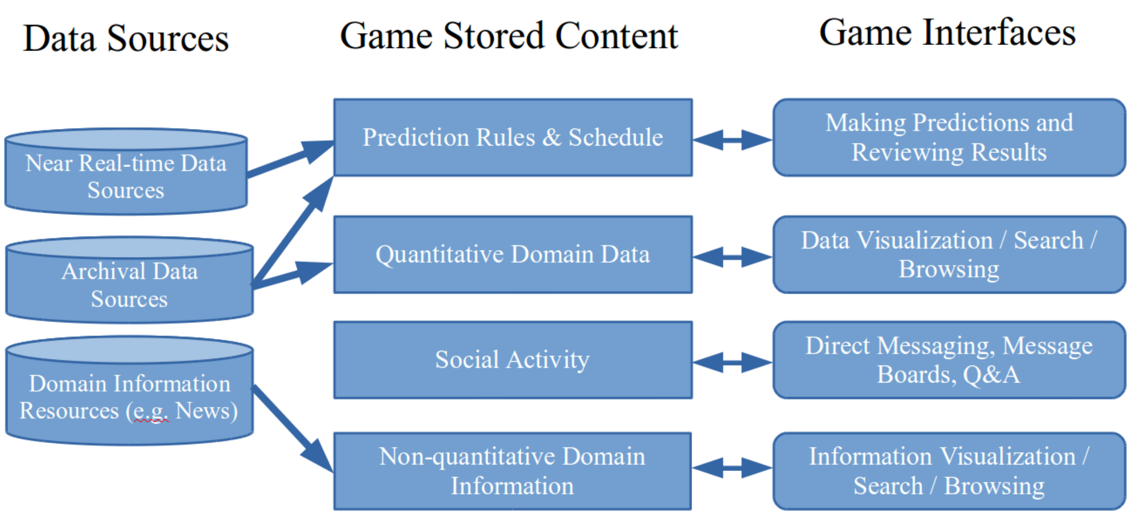

Figure 9. Primary data sources, stores and, interfaces

Figure 9 above illustrates the interactions among the three layers (columns) of our prediction game framework (from left to right). On the left are the data sources for domain data and domain-related information resources. Near real-time data sources and archival sources provide current and historical domain data respectively. In practice, they may be the same source. Domain information resources provide content for players to read to tie the prediction activity to real-world events. These resources may be news or social-network content such as related tweets or Facebook messages.

The middle column shows the content that needs to be stored as part of the prediction game server. The prediction rules and schedule are indispensable to the game mechanics. The prediction schedule dictates when the game starts, when it ends, and when the user predictions are due and scored. On a scoring date, the game engine uses prediction scoring rules to evaluate the user's submitted predictions against current domain data retrieved from the near real-time data source. We envision that the prediction rules and schedule are configured by the activity author who is setting up the prediction game activity. As a consequence, our framework requires data representations of prediction rules and schedules, rather than having them embedded (hard-coded) in the software code. The archival domain data is virtually static so may be retrieved once and reused across multiple prediction games where each game uses a subset of the domain data. For example, weather data can support both prediction games centered on temperatures and games centered on precipitation. Domain-related information resources on the other hand (e.g. domain-related news) depends on the prediction game. For example, a temperature-focus prediction game in the climate domain would filter climate-related news to select only that related to temperature. Lastly, the server needs to maintain the social activities of players during the game. The data stored contains the communications and interactions among players that are persistent over time (e.g. message board content, votes) along with any ephemeral communications not yet delivered.

The right column represents the client-side interfaces part of a prediction game. The prediction interfaces are required components of the game dynamics. They allow the players to make/update their predictions and view their scores once the predictions have been evaluated. This includes interfaces enabling the prediction activity designer (e.g. the instructor in the classroom example) to configure the prediction rules and schedule. The prediction-making interface is game-dependent. In the example of Fantasy Forecaster, players predict actual values of weather properties (left side of Figure 9) while an alternative game might have players ranking locations based on relative values (e.g. from hottest to coolest.) Interfaces enabling data interpretation present historical domain data as well as other domain-related information, such as news, social media content, images, etc. To support interpretation, these interfaces can include the ability for players to search, browse, filter, and visualize quantitative and non-quantitative content and include domain-specific visualizations. Lastly, fantasy sports foster community building through a variety of interfaces supporting social interactions among the players. These interactions may be a private conversation between two players (i.e. instant message or in-game email) or broadcast communications between all the players in a league or across multiple leagues such as unstructured discussion on a message board.

Appendix B describes the technical architecture of a prediction game system based on the aforementioned framework.

5. IMPLEMENTING PREDICTION GAMES: THE FANTASY CLIMATE CASE

Fantasy Climate is our first instantiation of a prediction game in the climate domain to include all the components described in our framework. It is based on weather data, building on our experience with Fantasy Forecaster, but includes the requirement for significantly more historic data in order to change the focus from weather to climate prediction.

Understanding climate requires considering long-term lower-amplitude variance in data that also includes higher-amplitude short-term variance (e.g. weather). Fantasy Climate encourages players to look for long-term change. Players are asked to make two selections from a given set of U.S. cities each scoring period: one for a location hotter than the historic norm and the second for a location cooler than in the past. These selections are scored on how much the observed daily high (low) temperature for the prediction date is above (below) the historical average for that day. Players quickly learn that it is easier to locate places getting hotter than it is to locate places getting colder and that warming is not uniform across the U.S. In the end, understanding both weather and climate effects is valuable for playing the game well.

We initially describe the major game components involved in prediction making and data analysis, including the main player interface, the interface for making predictions, and visualizations of historical data. These are common components of games using the fantasy sports model. The additional features identified as important from our formative efforts and included in the prediction game framework, the social activity and the non-quantitative domain information components, are described separately and are the focus of the evaluations of Fantasy Climate.

5.1. Fantasy Climate Core Components**

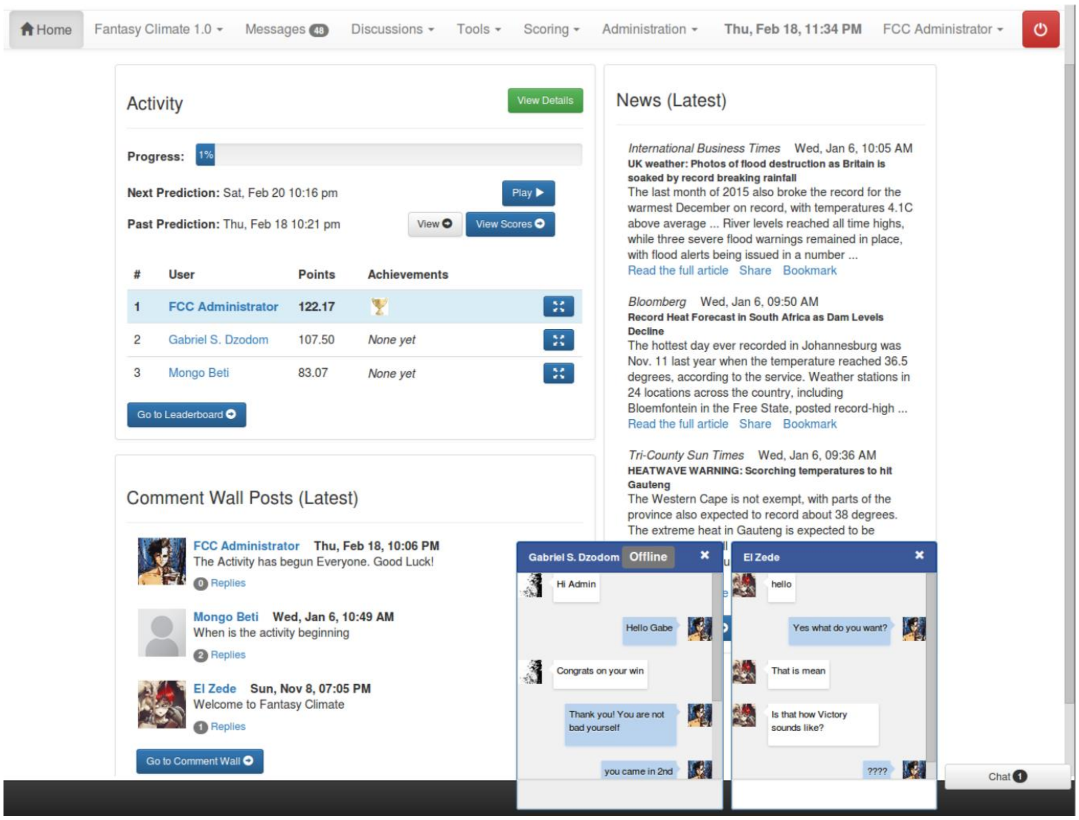

Figure 10. Fantasy Climate main interface

Figure 10 shows the entry point to Fantasy Climate; the main player interface. The Activity panel (upper left) shows the progress of the prediction activity. This includes the upcoming and recent prediction due dates and navigation to the prediction

---

** Part of this section is reprinted with permission from the paper: © 2016 ACM. Gabriel Dzodom and Frank Shipman. 2016. Data-driven Prediction Games. In proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA' 16). Association for Computing Machinery, New York, NY, USA, 1857 – 1864. DOI:https://doi.org/10.1145/2851581.2892546

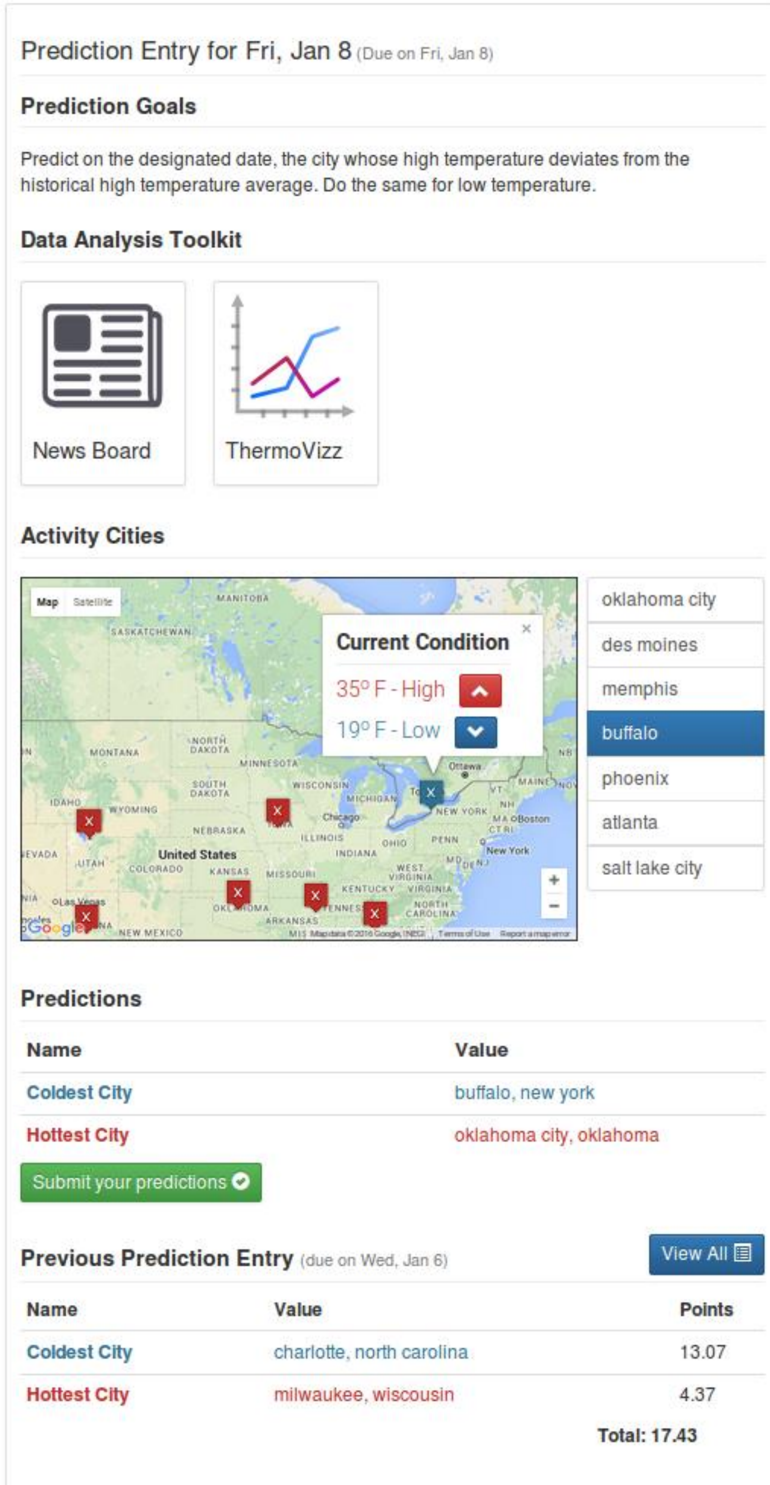

Figure 11. Fantasy Climate prediction interface

interface (Figure 11) along with detailed results from prior predictions. Below the dates is a snippet of the scoreboard incorporating the player's performance and the option to go to the full scoreboard. The main page also includes recent player communications from the comment wall and a few snippets of domain news. For more player communications or domain news, players can navigate to the full interfaces for these components. Thus, the main interface is a launching pad for players' interactions with the game.

To make, update, or view a prediction for the current round of the activity, the player navigates to the prediction interface shown in Figure 11. The top portion of this interface includes a description of the activity directives and links to tools available to aid in making predictions – the News Board and ThermoVizz. A map showing the cities part of the current round of the activity is shown in the center, with the player's current predictions shown below. Finally, the prediction interface contains a history of prior selections at the bottom. This includes the details of the recently scored prediction with a button that redirects the player to a view of all of their previously scored predictions.

Thermovizz (Figure 12), a data visualization tool to support prediction making, graphs the high (or low) temperatures for a specific date (the next scoring date) of the cities that are part of the current prediction round over the previous 50 years. Thermovizz enables players to compare temperature data between locations by toggling the visibility of data for each location. Thermovizz includes graphs of the yearly data (shown in Figure 12.) This graph requires players to interpret the variance in values from year to year. Thermovizz includes an alternative graph that presents the regression line for the yearly data that allows players to more directly compare the computed slopes making the climate change component of the prediction task more obvious, although weather plays a significant role in the scoring as frontal boundaries have a strong effect on observed temperatures.

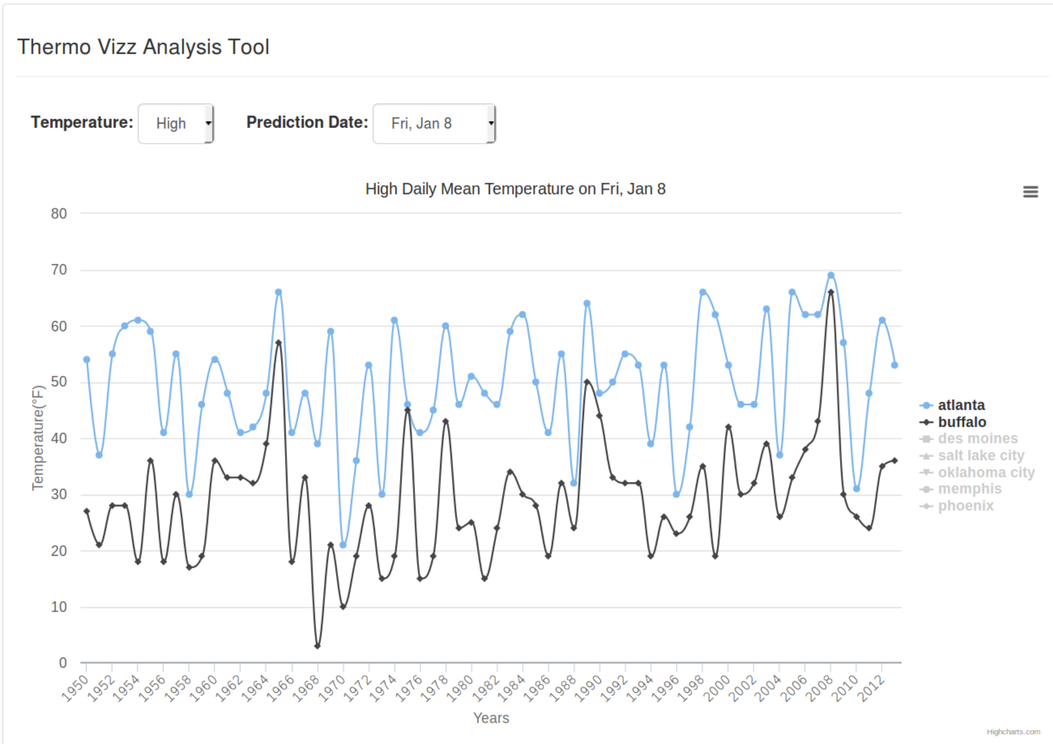

Figure 12. Thermovizz data analysis tool

Fantasy Climate promotes competition through an interactive leaderboard, a snippet of which is part of the main interface. Players can easily switch between an overall view of the scores, ranks and achievements of the players in the league to a detailed view of the scoring for a specific prediction round. There is also an achievement board showing any trophies earned (e.g. if they rank first in a prediction round.) It is common practice in online gaming to show achievements as badges or trophies to increase player motivation [2,40,41].

5.2. News in Fantasy Climate[††]

A finding from the survey of fantasy sports players was that news and online discussions played an important part of some players' decision-making processes. Two news providing interfaces were developed and fielded in the context of Fantasy Climate. NewsBoard, which is part of the main interface, provides a list (newest to oldest) of snippets of weather or climate related news. The second interface, GeoNews, presents news in a map interface based on the location(s) discussed in the articles.

To gather the news articles, Fantasy Climate requests articles from news provider sites using key phrases related to weather and climate (e.g. "record high temperature".) The metadata for the returned articles, the URL, date published, summary, title, author, are retained in a local database as part of the game server. For each news article, the Stanford's NLP Named Entity Recognizer is used to extract locations mentioned in the article and the frequently-mentioned locations are selected as the geographic focus of the article. Finally, when not available from the news source, a summary is generated and stored using a text summarization engine (e.g. TextSummarization.net.) The list of news articles is updated periodically to keep news current.

GeoNews (Figure 13) is a news interface that identifies the geographic focus of news on a map (openstreetmap.org). The news articles are grouped by their locations of focus. A marker on the map refers to the group of news articles that discuss that location. Clicking on the marker brings up an information window showing the title, a link to the

[††] Part of this section is reprinted from Effect of Visualization of News Articles in Data Driven Games, by Akshay Kulkani, Texas A&M master's thesis. http://hdl.handle.net/1969.1/156977. Copyright 2016 Akshay Kulkani.

full version and, the summary of the first article at that location. GeoNews also features a list view of the news articles adjacent to the map (on the right.) Each article in the list shows the title, associated location tags, and options to display the summary and navigate to the full article. Also, the list view is coupled with filters at the top-right and bottom on the viewport of the map. The filters at the top allow the user to select the time period for news displayed in the map and list. The location filters at the bottom are the locations associated with the current round in the Fantasy Climate game. Due to discussions about weather often being about larger geographic regions, players can toggle between seeing all available news and just those about the locations currently part of the game activity. As the user zooms, pans, and selects locations in the map, the list on the right is updated to present news related to the selected/viewed locations. The filters for time frame, locations, and search terms can be combined to further limit the news shown.

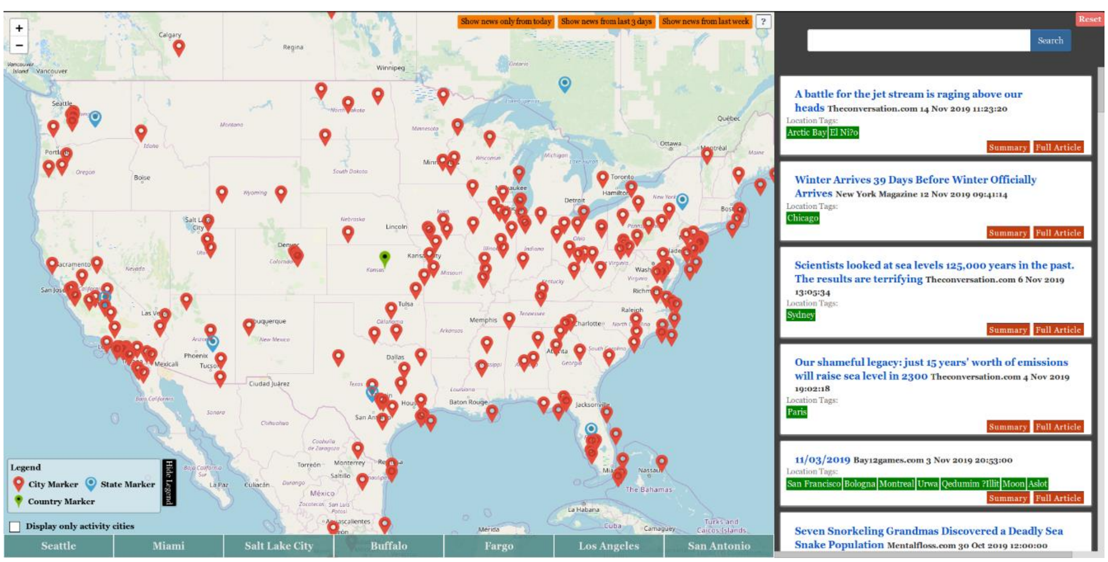

Figure 13. GeoNews interface includes map-based access and filters based on location, time frame and, search terms.

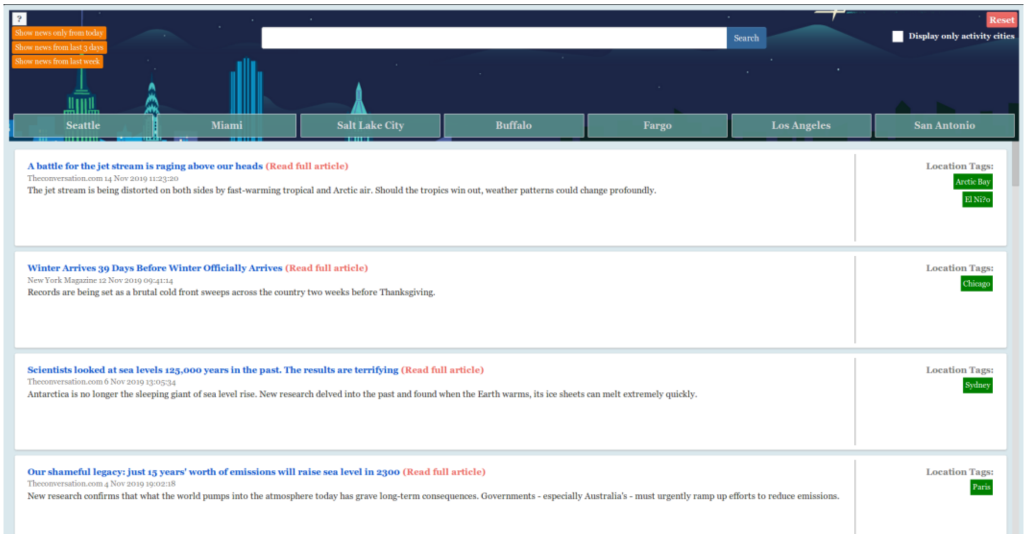

Figure 14. NewsBoard interface includes summaries and same filters as GeoNews

NewsBoard (Figure 14) is a news interface that presents the same news articles available in GeoNews in a more traditional list view. NewsBoard uses the same preprocessing and database of articles and has identical features (e.g. filtering based on time, location, and search terms) to GeoNews except for the map. As with many news lists (e.g. Google News), rather than having to ask to see the summary, NewsBoard presents the summaries by default. Players can again focus in on the locations that are part of the current round of activity and browse more generally via search.

5.3. Player Communication in Fantasy Climate‡‡

Fantasy Climate includes four ways for players to communicate with each other during a prediction activity. Two are for individual player-to-player communication and two are for open discussions between all players.

---

‡‡ Part of this section is reprinted from Asynchronous and Synchronous Communications' Effect on User Engagement in Prediction Games, by Meghanath Reddy Junnutula, Texas A&M master's thesis. http://hdl.handle.net/1969.1/155530. Copyright 2015 Meghanath Reddy Junnutula.

The two player-to-player messaging systems are in-game email and in-game chat for asynchronous and synchronous communication respectively. Two conversations using in-game chat are shown at the bottom of Figure 10. The email and chat interfaces are familiar for most players.

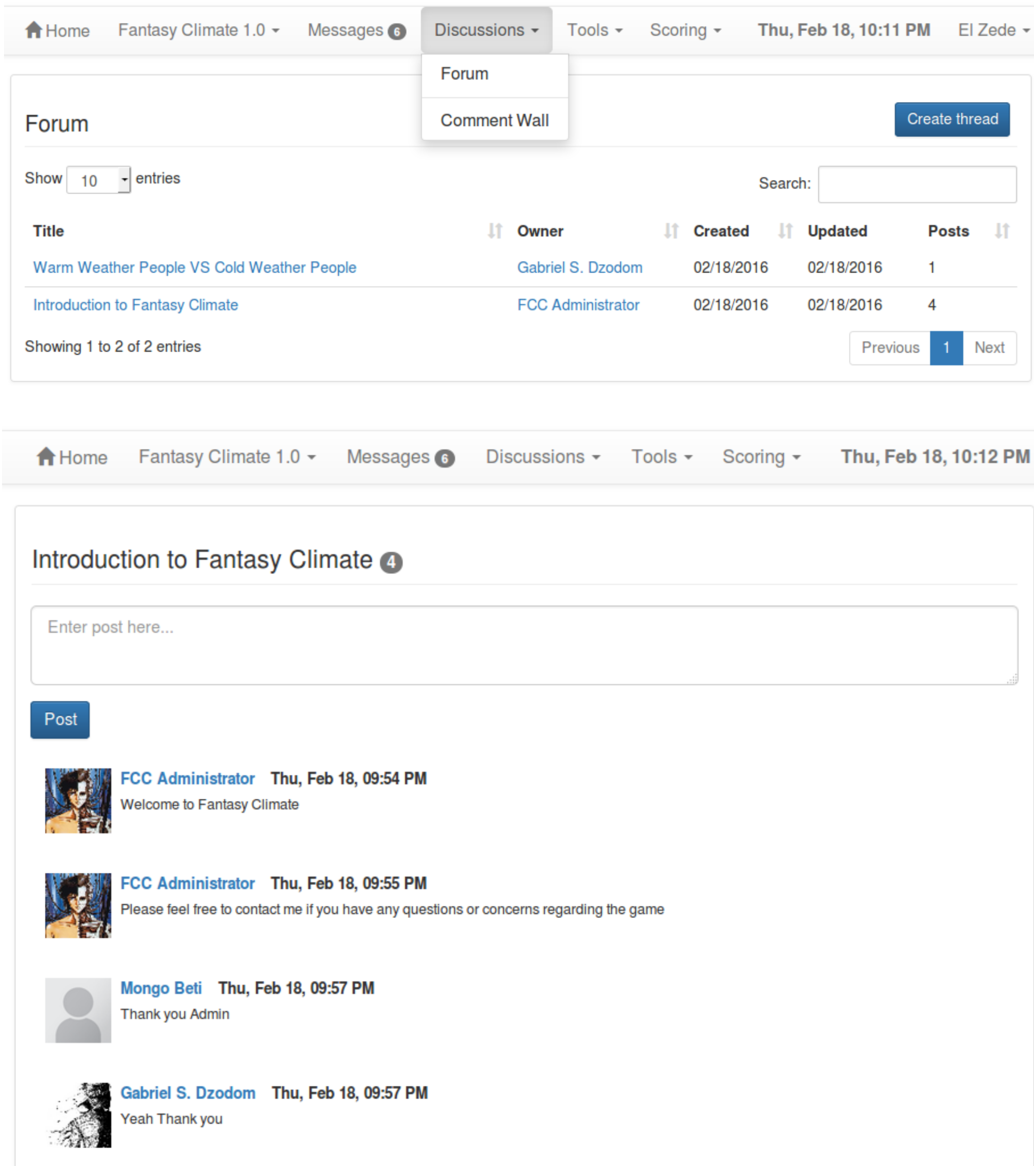

Figure 15. In-game forum interfaces. Top shows the list of topics while the bottom shows a specific discussion

The open discussion options are a comment wall, which is a list of public messages with replies ordered from newest to oldest, and a discussion forum, which provides an archive of messages on indicated topics. A portion of the comment wall is presented in the lower left of the main interface (Figure 10) while the forum interface is shown in Figure 15.) When navigating to the forum, a player is presented with the topics of on-going discussions. Players can navigate into a topic to read and contribute or create a new topic for discussion.

Appendix C describes Fantasy Precipitation, another implementation of our prediction games Framework. Fantasy Precipitation is again uses weather data but has players predict whether precipitation will or will not occur in given locations.

6. EVALUATING PREDICTION GAMES: THE FANTASY CLIMATE CASE

Fantasy sports are very popular in part because of widespread interest in their domain. The larger question of prediction games is, can this type of activity be engaging in non-entertainment domains. Although domain-independent, our prediction games framework incorporates many elements that make Fantasy Sports engaging such as competition, social interaction, presentation of domain news, and support for data analysis (e.g. visualization). Exploring the effect of such elements individually will provide guidance in the design of prediction games.

To begin answering the question of how best to increase engagement in non-entertainment domains, we have performed two evaluations of Fantasy Climate to examine the effects of: (1) the variety of player-to-player communication tools, and (2) the two domain news interfaces. Both studies focus on the engagement effects (e.g. do people play more) while also giving insight valuable to improve the design of prediction games. The study of communication tools also examined the amount and types of use of the different communication options while the study of alternative news interfaces additionally examined effects of news on the quality of player predictions.

The study of in-game communication tools and the study of domain news interfaces were conducted with slight variations to the version of Fantasy Climate already described. Here we describe aspects common across the two studies with study-specific details described in the later subsections.

Participants in both studies were recruited via word of mouth from the university community (mostly students) and had no prior knowledge of the prediction games project. They were provided a video tutorial covering the features and mechanics of the game. In addition, a 'Rules' button was available at the top of every interface that presented a description of the rules of the activity and how to play the game. Players accessed the game via a web browser on their own computers.

In both studies, the overall prediction activity lasted 14 days and included three prediction rounds. The first two days were a pre-activity stage to let the players familiarize themselves with the game goals and interface. During the next 9 days, a prediction submission was due every three days. Each round included 7 locations to learn about and select from, for a total of 21 unique U.S. locations. Each prediction was for the day after the submission due date, which was also when predictions were scored. The final two days of the activity allowed participants to review their scores and to socialize.

For each round of the activity, players were given a set of US locations and made two predictions: one for the warmest location compared to historic norms and the other for the coolest location compared to historic norms. The following formula was used to score the players' predictions:

Score for warming location = 50 + Observed high – Historical average high

Score for cooling location = 50 + Historical average low – Observed low

Weather data for these studies (both historical and current) came from Weather Underground (wunderground.com). The historical data visualized in Thermovizz was from the year 1950 to 2013. We defined the high/low historical averages as the mean of high/low temperatures for the first 30 years of this data (01/01/1950 – 12/31/1979).

6.1. Effects of Communication in Prediction Games[§§]

Prior work (section 2.1) and our survey (section 3.1) show that social interaction is a motivational factor in fantasy sports. To further understand this effect and how alternative communication channels are used by players, a user study of Fantasy Climate was designed to answer the following question:

How do different in-game communication tools affect engagement with the activity?

In particular, the study aims to gain insight into the relative value of in-game email, chat, the topic-based forum, and temporally-structured comment wall through measures and observations of their use.

6.1.1. Experimental Setup

We recruited 27 participants from around our university campus for the study. There were 17 males and 10 females in the age range of 18 - 35 years. The majority were studying engineering. All had considerable experience at operating a computer and accessing websites and web applications. The participants were randomly split into two groups: Group-1 (14 participants) and Group-2 (13 participants). Then we varied the mode of communication (asynchronous or synchronous) that each group could use over the period of the activity as shown in Table 1. During the pre-activity stage, no communication channels were available to either group. During the first prediction round, Group-1 could only employ asynchronous communication channels (i.e. email,

[§§] Part of this section is reprinted from Asynchronous and Synchronous Communications' Effect on User Engagement in Prediction Games, by Meghanath Reddy Junnutula, Texas A&M master's thesis. http://hdl.handle.net/1969.1/155530. Copyright 2015 Meghanath Reddy Junnutula.

comment wall, and forum) and Group-2 only synchronous channels (i.e. chat.) During the second prediction round, the configuration was reversed for the groups. For the last prediction round through the end of the post activity phase, all communication channels were enabled for both groups. Web analytics was used to log users' interactions during the activity and a questionnaire was administered at the end of the activity.

Table 1. Communication tools available to the two groups during game (PRE=pre-activity, POST=post-activity, PR =prediction round).

	PRE	**PR1**	**PR2**	**PR3**	**POST**
Group – 1	None	Async	Sync	Both	Both
Group – 2	None	Sync	Async	Both	Both

6.1.2. Results and Discussion

The in-game communication methods were compared via player reports of about each communication tool, quantitative data about the amount of use each tool received, and an analysis of the types of communication found in each of the available forms.

What were the preferred ways of communicating? A post-questionnaire asked players about their experience with Fantasy Climate and the communication tools. Players were asked which communication tools kept them involved in the game. The forum had the most positive responses (24 of 27). After the forum was the comment wall (19 of 27), email/direct messages (11 of 27) and finally instant messaging/chat (7 of 27.) These results are shown in Figure 16. A related set of Likert-scale questions asked whether each communication tool caused players to play longer than intended – an effect common for entertainment-focused games and relevant to increased engagement. The results (Figure 17) confirmed the prior reports. The forum again ranked first with 63% agreeing that they played longer because of it, followed by 59% for the comment wall, 48% for email/direct messages, and 40% for instant messaging/chat.

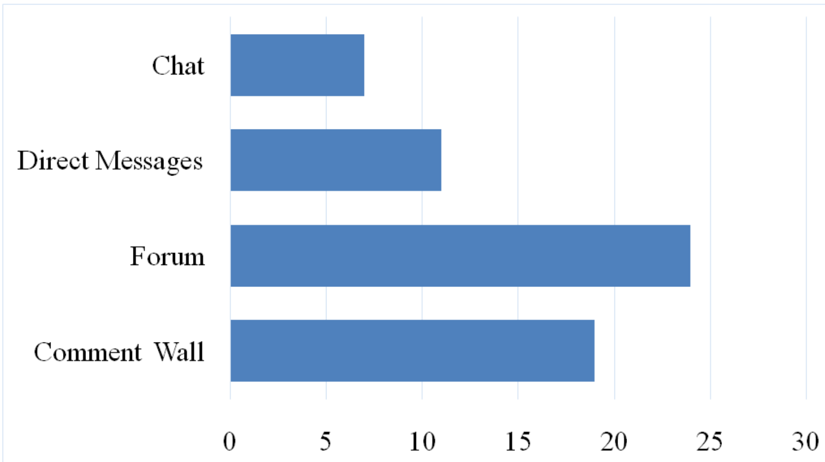

Figure 16. # of players (out of 27) indicating that communication tool kept them engaged

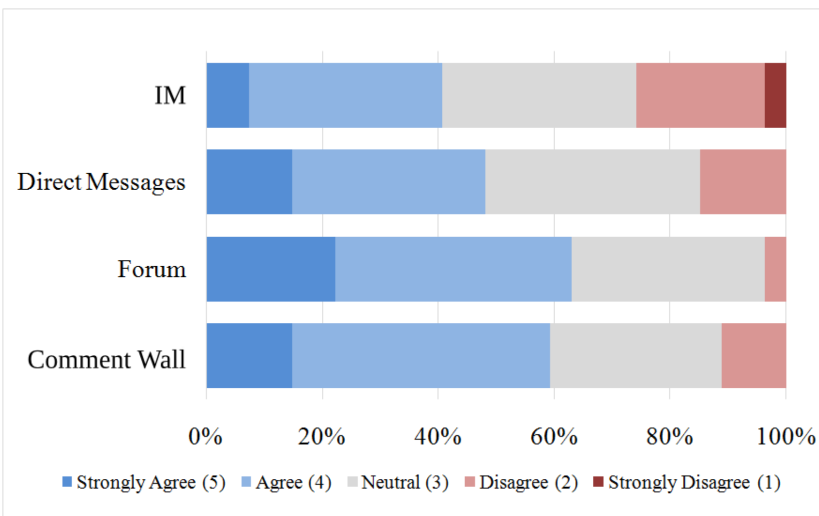

Figure 17. Player sentiment on whether the tools caused them to play longer than intended

Open-ended questions encouraged the participants to elaborate on the usefulness or lack thereof of each of the communication tools. Information exchange was the main theme that emergence in their responses on the forum (92% of responses): "*...the discussions in the forum helped us to think of other ways to predict the data using different statistical methods.*" Furthermore, the responses describing the forum contained phrases like "*most important*", "*effective*", "*one of the highlights*", "*best part*", "*best way*" or "*most essential*".

The same information exchange theme was dominant in the comments regarding the comment wall and instant messaging (84% of responses for comment wall, 77% for IM): "*Yes, it [comment wall] did keep me engaged in the game. It helped in communicating with different people resulting in exchange of different views and perspectives regarding prediction " and "...I believe that the instant messaging system was very useful in exchanging ideas with other and getting inputs from them too. These kind of instant messaging communicating systems must be included for effective performance of a player.*"

The views on the email/direct messages were nuanced. For instance, 23% of the participants explained that direct messages were an unnecessary communication tool: "*...the forum is good enough to exchange ideas. The direct message feature is an additional tool that distracts the players.*" Others (42%) reported that they used directed messages to exchange information with selected other players (15%). Finally, Fantasy Climate sends reminder notifications via direct message to players as the prediction deadline approaches. Some responses (19%) highlighted this feature as a useful role for this mode of communication.

Finally, the instant messaging application, the only synchronous communication tool, had the least effect on engagement according the evidence from analytics and player feedback. The following comment from one participant about their lack of use of IM may indicate the principal reasons: "*It was not common for users to be online at the same time. Also trying to initiate instant messaging with unknown users felt awkward.*" Not only did our participants not know each other, the league size was too small for a reasonable of number of users to be online most of the time.

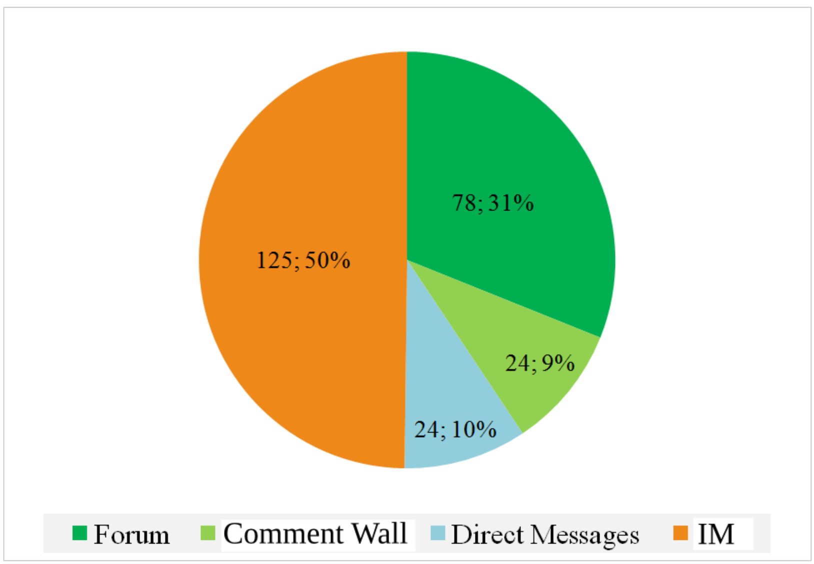

Figure 18. # of messages per communication tool

How much communication was really happening? Another measure of the effect of the communication tools is to examine how much they were used. For our analysis, we removed all communication involving the activity administrator as our focus was interactions among the players. Figure 18 shows the number of messages between players in each of the four tools. Of the 251 total messages, the number of messages in the instant messaging tool was equivalent to the number of messages from all the asynchronous messaging mechanisms combined. Figure 19 also shows that the forum was used for more than 60% of the asynchronous messages.

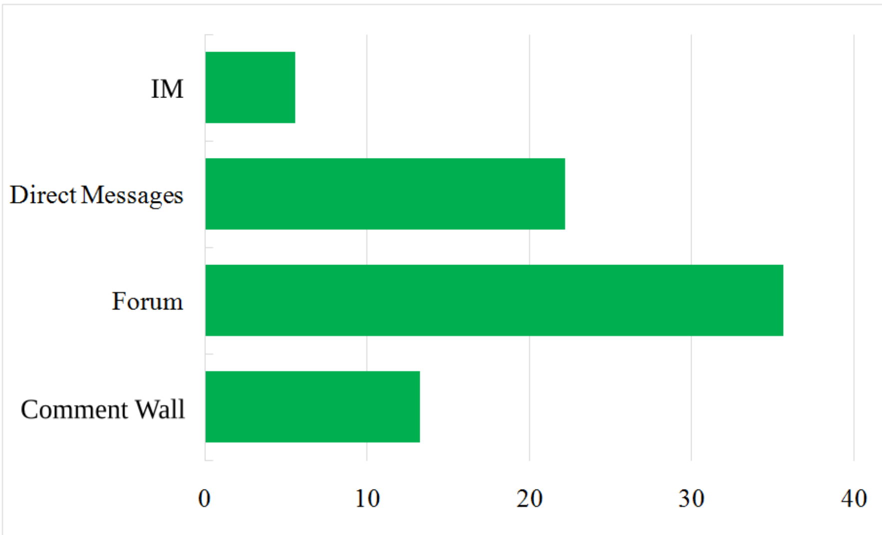

Figure 19. # of words per message

One reason for the high number of messages in instant messaging was that such messages were necessarily part of a real-time conversation that includes separate messages for salutations, questions, responses, etc. while messages in the forum, comment wall, and direct messaging system (in-game email) often concatenated these components into longer messages. Figure 19 shows the average number of words per message across the four modes. Forum posts were the longest (35 words on average) while direct messages (22) and comment wall posts (13) were still well above the average of 6 words per IM message. Figure 20 shows the total number of words across all messages for each option. This shows that 64% of the words across all conditions were found in the forum.

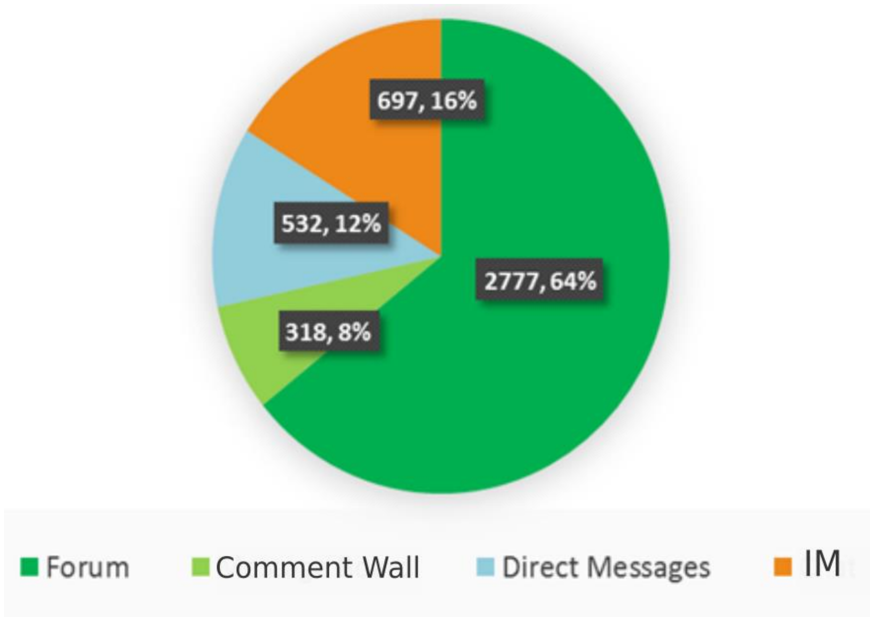

Figure 20. # of words per communication tool. Reprinted from [31]

What were players talking about? Knowing what type of communication happens with each tool can help understand player activity. We open coded the content of each of the 251 messages to understand the type of activities that were driving social interactions among players in each of the communication tools. The four top-level categories that emerged were, in the order of their frequency across all communication tools: (1) making and scoring predictions (116 messages or 46%), (2) greetings (63, 25%), (3) system-related questions (46, 18%), and (4) trash talking/casual conversations (26, 10%).

Discussions regarding prediction making dominated the discourse in the forum over the life of the activity. 91% of the forum posts were in this category. The contents of the forum posts were often about data analysis techniques and tools: "*Since we have time series data we can use time-series forecasting in R to predict the future values*", sharing of resources: "*@hoshi http://www.weather.gov/ This is the best way!!*", or requests for help or advice: "*OMG! This is getting too technical...Differencing, logarithms to stabilize variance...!!Is there any simple approach which can be comprehended by someone who is out of touch of Statistics for quite a while??*" Although substantially fewer, similar discussions of prediction techniques also occurred on the comment wall (54% of wall posts) and occasionally in instant messaging conversations (16% of messages).

When messages were not related to predictions, other modes of communication were more common. The comment wall was used for rapid questions and announcements. Participants mostly employed instant messaging and direct messages to support each other: "*Just go to the scoreboard and select overall on the filter and show 25 people on the show rows...*" Finally, casual and trash talking were common on the chat, but were unlikely on the asynchronous communication system.

Overall, the results provide evidence that just like in Fantasy Sports, social interaction is an important part of the gameplay in Fantasy Climate. Players reported that the asynchronous tools had a more positive effect on engagement than did IM. The forum was favored by the majority of players and was the hub for many interactions among members. Forum topics often involved in-depth discussions of prediction techniques or data analysis methods. The comment wall fulfilled a similar role as the forum but was less used. Direct messages were valued as a way to receive reminders about activity due dates. Finally, IM was used for casual conversations and trash talking but required multiple players to be online simultaneously and to overcome the initial hurdle of initiating discussions.

6.2. Effects of News Presentation in Prediction Games***

Surveys show that a variety of information resources, including news, play a valuable role in the decision making process of fantasy sports users [17]. In addition, the in-game presentation of domain-oriented news may help to engage players by building stronger connections between what is happening in the game and what is happening in the real world.

To examine the effect of in-game news on decision making and engagement, we explored how players performed with the two news visualizations from Section 2.4.2, GeoNews and NewsBoard. Fantasy Climate provides weather and climate related news to help the players make good predictions. However, the large quantity of news articles may be an obstacle in the data gathering and analysis activities of players. This study looks at the effects of how news is presented on measures of user engagement and prediction accuracy.

6.2.1. Experimental Setup

We recruited 29 participants (22 males and 7 females) from around our university campus for the study. 27 of them belonged to the age group of 23-27 and 2 participants were in the age group of 18-22. Most of them were studying engineering.

*** Part of this section is reprinted from Effect of Visualization of News Articles in Data Driven Games, by Akshay Kulkani, Texas A&M master's thesis. http://hdl.handle.net/1969.1/156977. Copyright 2016 Akshay Kulkani.

All had considerable experience at operating a computer, accessing websites, and using web applications.

The participants were randomly assigned to two independent leagues with identical activities except for the news presentation interface. GeoNews was used to present news in the first league (14 participants) and NewsBoard in the second (15 participants.) The two leagues were run concurrently with the same subsets of locations available in each round meaning the difficulty of the prediction tasks was identical across the two groups. Web analytics were used to log participants' interactions and a post-activity questionnaire collected participant experiences and perceptions using a combination of Likert-scale and open-ended questions.

6.2.2. Results and Discussion

The effects of in-game news presentation were compared via player feedback from the post-activity questionnaire, quantitative data about the time spent in the game and the amount of use of the news interfaces, and an analysis of the quality of predictions in the two leagues.

How did the two news interfaces affect player perceptions? Participants in both groups were positive about the availability of in-game weather/climate news. 71% in the GeoNews league (Group 1) and 66% in the NewsBoard league (Group 2) at least agreed with "*Exploring the news articles kept you engrossed in the game*". Similarly, 78% in GeoNews users and 73% of NewsBoard users agreed to some degree with the statement "*Exploring the news articles was part of what made the game fun.*" Thus, there was little difference in the two groups' reactions to the availability of news.

Where perceptions differed between the groups was when they were asked about their engagement with the game as a whole. First, shown in Figure 21, more than half (56%) of GeoNews league members at least agreed that they played Fantasy Climate longer than they meant to compare to only one-fifth (20%) of NewsBoard league members. Similarly, when asked whether they really got involved in the game, 85% of those in the GeoNews group at least agreed compared to only 53% of those in the NewsBoard group (Figure 22.)

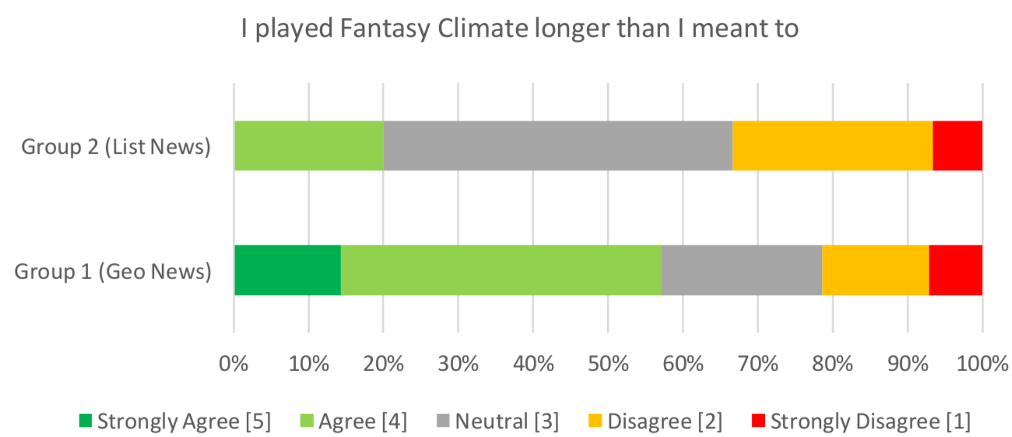

Figure 21. Players' perception of playing more than intended. Reprinted from [34]

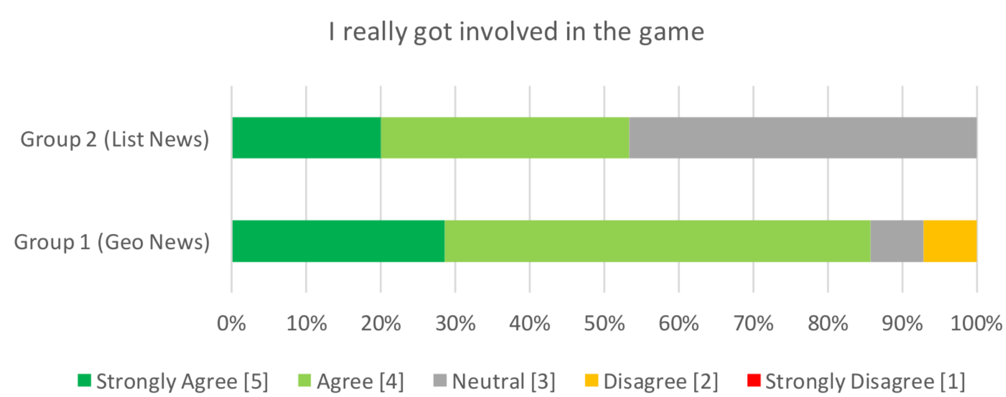

Figure 22. Players' report of feeling involved with the game. Reprinted from [34]

How did the two news interfaces affect time spent with news and with the game? We collected interaction data using Google Web Analytics [15,44,58]. The focus was on the following metrics as a way of gauging engagement: page views, number of sessions, average session duration, and average overall time spent. The results across all of these metrics indicate a consistent pattern at the overall game level and at the news interface level. The GeoNews users had a higher traffic than did the NewsBoard users. In particular, they had 32% more unique page views, 23% more web sessions, and 68% more unique page views in the news interface provided.

But page view metrics can end up high due to poor navigational cues resulting in users visiting many pages only to realize they did not get what they intended (or navigated to news that was not interesting.) Session duration and time spend in the game are more direct indicators of player engagement.

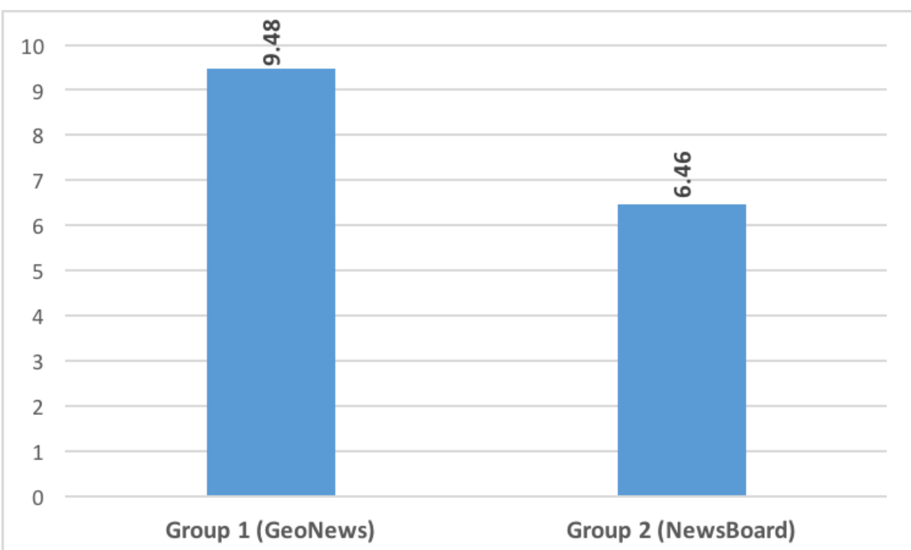

Figure 23. Average session duration (in minutes.) Reprinted from [34]

As seen in Figure 23, not only did the GeoNews group visit the game more but its members spent more time per visit. The average session duration of GeoNews players was 9 minutes and 29 seconds compared to 6 minutes and 28 seconds for the players using NewsBoard. When looking at the amount of time participants spent playing the game through the entire activity, players in the GeoNews group were active in the game for more than 115 minutes across the two weeks compared with slightly less than 60 minutes of activity for players in the NewsBoard group (right portion of Figure 24.) In terms of time spent in the news interface, GeoNews users logged a bit over 34 minutes compared to the less than 12 minutes for those with the NewsBoard interface (middle of Figure 24.) Finally, with regard to the level of engagement with individual news stories located, the GeoNews group spent an average of 4 minutes and 44 seconds per page view compared to 3 minutes and 15 seconds for NewsBoard (left side of Figure 24.)

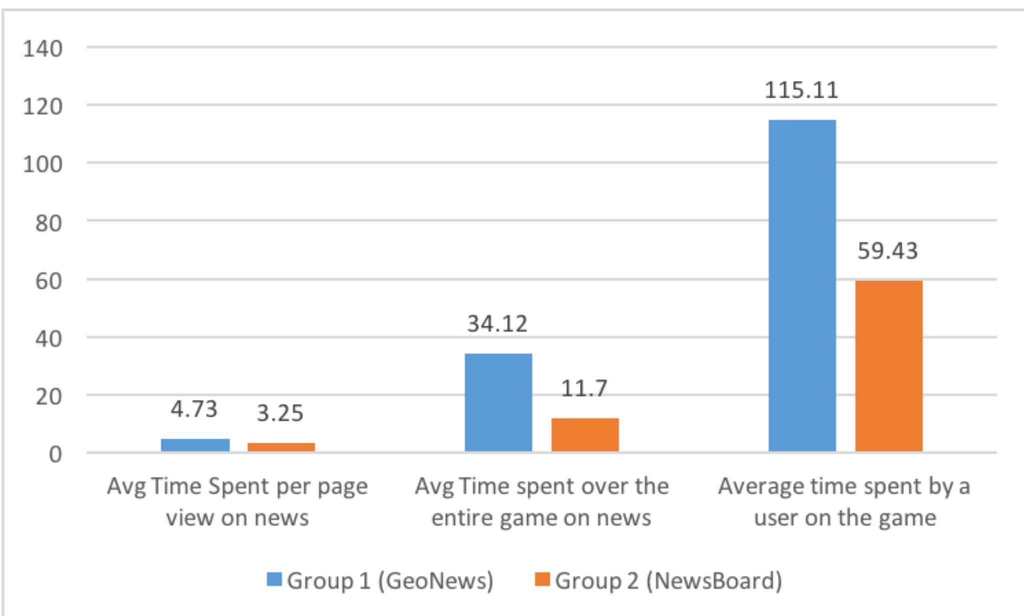

Figure 24. Average time spent on the game, on news, and per page view (in minutes.) Reprinted from [34]

In general, the analytics indicate that GeoNews participants interacted more often and longer with Fantasy Climate than NewsBoard participants. This was also true when considering only the news interfaces. We could potentially interpret the difference as players using external sources to find additional information in the NewsBoard case, but not all of the extra time spent over the course of the league was in the news interface. In fact, only about 22 of the 55 extra minutes (on average) over the activity was in the news interface.

Does the news interface have an effect on the quality of player predictions?

The above results provide evidence that the news presentation influences player engagement with Fantasy Climate. However, this outcome says nothing about the effects on player performance in the game activity. Does the extra time result in better predictions?

To compute the players' performance in each prediction round, we ranked the round's cities based on the deviation of their high/low temperature from the historic average. Each city had two ranks, the warming rank and the cooling rank. Rank 1 meant that the city was the best of the cities available for selection; rank 7 was the worst. A principal advantage of our ranking is that a cooling rank 1 meant that the city was the best available option for the round even if the observed low temperature was actually warmer than the historic average low. Thus, a player's performance in a prediction round was the average of the ranks of his predicted cities.

A lower score is better when considering the average rank of the predictions. Overall, the average performance over the entire game of the players in the GeoNews league was better (3.28) than the one (3.96) of the players in the NewsBoard league. This difference is statistically significant p=0.015 (two-tailed t-test.)

This result also held when the performance was broken down by prediction type as shown in Figure 25. Also in this figure, it is interesting that the warming location predictions were more accurate than the cooling location predictions for both groups and there was a relatively consistent gap between performance on warming predictions and cooling predictions. The reason may be that cool predictions are more difficult to make because low temperatures are rising. This is the kind of insight we expected our players to develop as they play Fantasy Climate over a period of time.

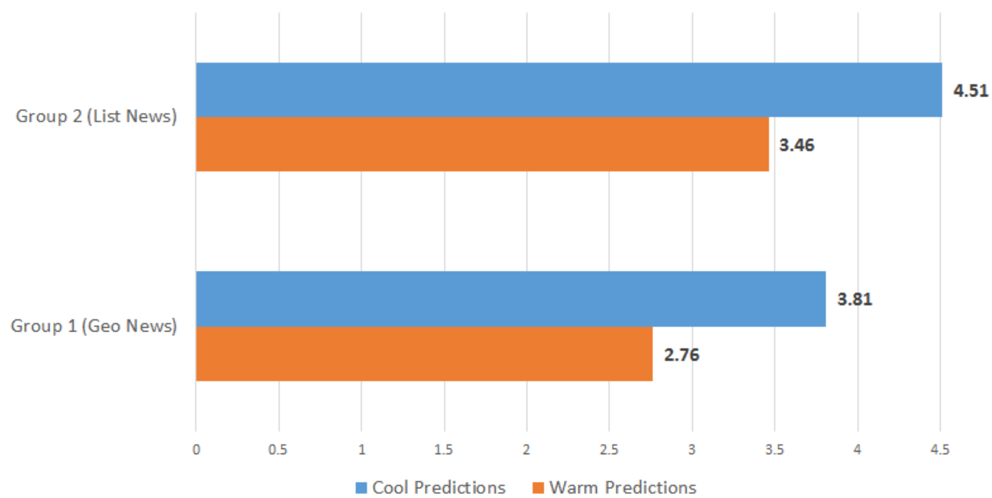

Figure 25. Overall average performance per prediction type (lower is better.) Reprinted from [34]

Finally, how do player predictions change over the course of an activity? Our hope is that, by engaging with data and news about the domain, players develop a better mental model over time and so their predictions should improve over time. Unfortunately, Figure 26 shows this is not always the case. It shows that the predictions from the GeoNews group and the NewsBoard group were similar for the first round of the activity. But, while the GeoNews group predictions improved over time, the NewsBoard group predictions worsened. A possibility in activities like this is that engagement early on due to novelty will subside resulting in decreased prediction quality. Features that increase engagement, including effective news presentations and player communication tools, are meant to reduce such drop-offs in effort.

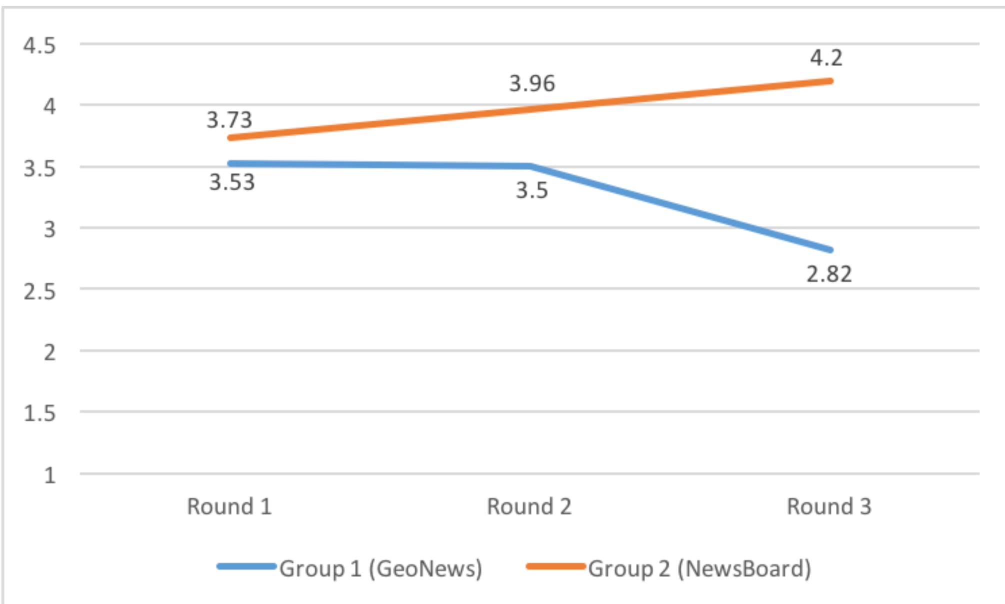

Figure 26. Average performance per prediction round. Reprinted from [34]

In Summary, the experiment results indicate that news presentation influenced players' performance in Fantasy Climate. The results suggest that not only the information available, but also its presentation, helped players make better predictions over time. From the web analytics, we observed that news presentation increased players' activity in Fantasy Climate. Participants in the GeoNews group visited the game and the news interface more and longer than their NewsBoard counterparts. Taken together, it appears that the GeoNews interface engaged players more and the extra time they spent in the game improved the participants' understanding of the domain, resulting in better predictions.

7. AUTHORING PREDICTION GAME ACTIVITIES

Given the potential for prediction games to increase engagement with data and other content, we envision prediction games in formal and informal educational environments. In such settings, it is valuable to tailor the prediction activities to the particular learning goals of the setting and the background knowledge of the participants. Tailoring existing and creating new activities implies enabling the authoring of prediction game activities without requiring programming. This goal led to the exploration of authoring support.

Inspired by work on domain-oriented design environments and instructional design environments, we have designed and implemented an authoring environment to create and customize prediction activities for a prediction game without computer programming. The particular environment designed and evaluated is a wizard user interface that guides the user through a series of steps in order to author the prediction activity. The following sections describe the development process of the authoring environment for prediction games starting with related work on environments supporting design. After this is a discussion of the prediction activity specifications followed by their implementation as an authoring environment for tailoring Fantasy Climate activities. Finally, we report on an evaluation of the authoring environment to determine how different components of such systems affect the authoring task and, to discover the kinds of prediction games that people may find engaging.

7.1. Example Scenario

In one example scenario, an instructor whose class is learning about precipitation patterns in coastal U.S. locations may create and customize a prediction activity for Fantasy Climate where players will make two prediction for a specific date and a given set of locations: (1) the wettest location, the one with highest precipitation accumulation on the given date, and (2), the least wet location, the one with the lowest precipitation accumulation on the given date. The customization includes building of the prediction schedule which specifies when student predictions are due and when they are scored; for example, predictions might be made weekly, due on Friday evening and scored on Sunday evening. Another customization is the definition of scoring rules for the predictions; for example:

Score for wettest location = Observed precipitation accumulation in inches

Score for least wet location = 5 – Observed precipitation accumulation in inches

At the end of her customizations, the teacher invites her/his students to join the game and participate in the new prediction activity.

7.2. Environments Supporting Design

According to Gerhard Fischer, domain-oriented design environments (DODEs) are computational environments that support activities involving a design process (e.g. software engineering) within a particular domain [20,21]. He proposed a domain-independent architecture for building DODEs with the following components: (a) a construction kit for modeling, (b) an argumentation component as a repository for the explanation of issues, answers, and rationales in the design domain, (c) a catalog as a collection of prestored designs or templates, (d) a specification component that allow for the description of design characteristics and, (e) a simulation component to simulate the artifact being designed [21]. Together, these components are meant to enable and proactively support non-programmer domain experts to create, reason about, and evolve new designs.

This architecture is well suited for the design of artifacts whether physical (e.g. circuit board) or digital (e.g. user interface). That said, not all DODE components are equally appropriate when considering the design of prediction games. Our proposed authoring environment primarily needs to support the definition of and management of the prediction activity. As a consequence, a component like the construction kit, which might be used to change the layout of the components in the player interface, is less important in our context while the catalog, providing tailorable examples, and the specification component, where the schedule of activity might be defined, are more necessary. The DODE architecture, with its emphasis on supporting non-programmers, is a source of inspiration for the design of the prediction activity authoring environment.

Closer to the context of our proposed project, there is a variety of authoring environments available which support instructional design. Some focus on providing instructors with tools for the production of course content. For example, interfaces that enable the integration of multimedia and interactivity into the course content [43]. Others are more complete by including features that let the instructor configure the learning experience [1,27,33]. For example, in Web-CT (now Blackboard), the instructor can decide how the course content is viewed and navigated, or whether to permit social interaction through the bulletin board or the chat [27]. In the prediction activity authoring process, the instructor will not be producing the content for students (e.g. the data or news stories) beyond descriptions of the activity. Instead the teacher defines the prediction activity and customizes how the students will experience the activity. Thus, environments like Web-CT that afford customization serve as models for the design of our prediction activity authoring environment.

8. PREDICTION ACTIVITY SPECIFICATIONS

A prediction activity is a custom instance of a prediction game. Hence a prediction game can yield several prediction activities (sometimes running concurrently). Creating and customizing a prediction activity will vary depending on the prediction game or the data domain. Furthermore, a prediction activity is similar to a league in fantasy sports. It has an author who creates and administers the activity and members who play by making predictions. They correspond respectively to the league commissioner and league members in fantasy sports. In a formal learning settings, the activity author may be the instructor and the activity members the students. Using the prediction games framework described in section 2.3 and the implementation described in section 2.4, we derive the following activity specifications that are customizable and required to author a prediction game: the activity name and directives, the prediction schedule, the pivot set and selection sets, the scoring rules, the tools supporting prediction making, the components facilitating social interactions, and the activity members.

8.1. The Activity Name and Directives

The name may be a keyword or a keyphrase that references the prediction activity like the title of a game (e.g. 'Extreme Locations League'). Since players will be making predictions on a schedule (e.g. once a week), the directives may serve as a reminder of the prediction activity objectives. As a consequence, the directives are displayed on the player's prediction interface (Figure 11). Here is an example for the prediction activity used in the evaluation of Fantasy Climate:

On the designated date, predict the city whose high temperature is warmest relative to its historical high temperature average. Similarly, also predict the city whose low temperature is coolest relative to its historical low temperature average.

8.2. The Prediction Schedule

The prediction activity schedule comprises a start date, an end date, and the players' predictions schedule. The start date is the date and time when the prediction activity becomes active. The prediction activity expires on the end date and becomes inactive. The active or inactive state of the prediction activity has consequence on the prediction game. For example in Fantasy Climate, when players who are not part an active activity sign in, the game user interface is minimal allowing them to only update their account profile settings (Figure 27.) Once the activity becomes active, the game interface is updated to reflect the activity customization. For example, new menu options appear on the navigation bar enabling the player to submit their prediction (Figure 10).

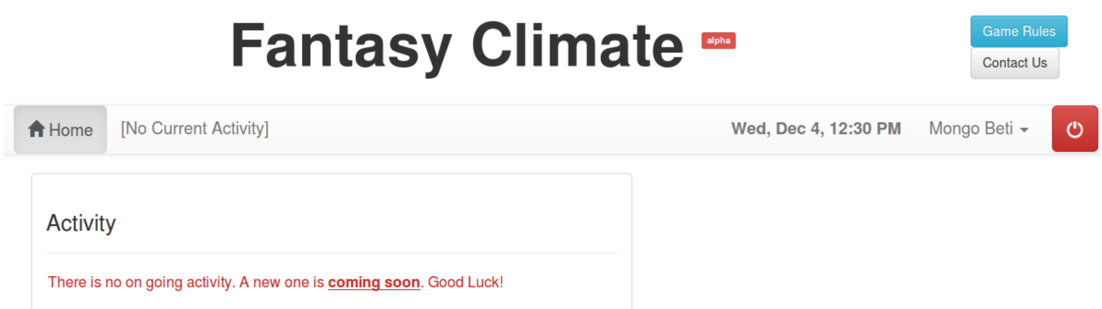

Figure 27. Main interface when there is no active activity

The players' predictions schedule defines the series of prediction rounds within the activity start date and activity end date. A prediction round consists of a due date and a scoring date. A due date is a deadline by which players must submit their predictions for the prediction round. Hence after the due date, players will not be able to update nor edit their predictions. The scoring date is when the players' predictions submitted for the prediction round are scored. Prediction rounds need not be sequential. The prediction game environment allows overlap between prediction rounds. This means the scoring date of a prior prediction round may be after the due date of the next prediction round. For example, an activity might have players make predictions daily but score them a week later as observations become available. As a consequence, after the due date of a prediction round, players can submit new predictions for later prediction rounds even though predictions for prior prediction rounds have not been scored.

One way the prediction game environment tries to draw back the players into the game is by sending them notifications about their predictions. There are two main notifications the players receive:

- the prediction submission reminders – which are messages reminding them to submit their predictions before the prediction round due date, and
- the scoring notification – which alerts the players that their predictions have been scored and that the results have been posted

The author of a prediction activity may configure the prediction submission reminders. The reminders are not required for the prediction activity to work and might not be wanted in contexts with regular in-person discussions of the activity (e.g. classrooms.) In such a case, the activity administrator can make it so the players will receive no notifications about the prediction round deadlines.

8.3. The Pivot Set and Selection Sets

The set of information available to players of prediction games can be considered open ended as players may seek out information not directly provided by or used for scoring by the prediction game. For example, in the survey of fantasy sports players, there were reports of considering whether an athlete was in the final year of his/her contract when determining who to draft.

When authoring a prediction activity, the domain data ingested by the environment necessarily limits the data that can be used to score or directly support prediction making. In the case of Fantasy Climate, the domain data ingested includes the full weather data for a number of locations. The activity author defines pivot sets and selection sets to bind the quantity of data involved in a particular prediction activity.

In the prediction game environment, we define the pivot set as the dataset through which the players interface with the domain data. In fantasy sports, the domain data is the athlete statistics and the pivot set is the set of athletes from which the fantasy sports players will be making selections to constitute their fantasy team. In Fantasy Climate, the domain data is weather data collected and the pivot set comprises the locations (or cities) for which weather data are collected. Thus, the pivot set is dependent on the prediction game and the domain.

A selection set is a subset of the pivot set. It is a set around which players' predictions will revolve. In a typical prediction activity of Fantasy Climate, players predict by picking one or more locations from a selection set that meet the activity objectives. The activity author creates and customizes the selection sets. To make the gameplay less repetitive, the prediction game environment facilitates the creation of multiple selection sets for a single prediction activity. Then, the activity cycles through the created sets so that in each prediction round the players will be making predictions from a different set of locations, assuming the activity author generated as many selection sets as there are prediction rounds in the activity. For example, in a prediction activity for Fantasy Climate where players are encouraged to observed long term temperature changes in U.S coastal cities, the selection sets will contain a variety of U.S coastal cities configured by the activity author.

8.4. The Scoring Rules

The scoring rules of a prediction activity define how the players' predictions are scored. In our prediction game environment, scoring rules are expressed as mathematical formulas. For example, let us remember the prediction activity for Fantasy Climate that we have used for the user studies described in chapter 6. For this activity, the players are asked to make two predictions given a set of locations: (a) one for the warmest location compared to historic norms, and (b) the other for the coolest compared to historic norms. Scoring rules for this prediction activity may be:

Score for warming location = 50 + Observed high – Historical average high

Score for coolest location = 50 + Historical average low – Observed low

The scoring rules are central to the prediction activity as they define how the players compete and affect the focus of players as they play the game. In the aforementioned example, the activity author may discover (based on the trends in the data) that predicting the coolest location is more challenging than predicting the warmest location. Thus, she may update or redesign the scoring rules to increase the reward for the more challenging prediction. Her new scoring rules may look like the following:

Score for warming location = 50 + Observed high – Historical average high

*Score for coolest location = 50 + 2 * (Historical average low – Observed low)*

By doubling the difference between the historical average low and the observed low, the second in the formula, it is more valuable to players to spend time making good predictions for the cooling location than it is for the warming location.

8.5. Supporting Prediction Making

Prediction games may provide built-in data analysis tools to support the players in their prediction making. In Fantasy Climate, the Thermovizz tool (Figure 12) visualizes high (and low) temperatures over time, helping players to discover long term trends in the data. While the availability and customization of data analysis tools is not required for a prediction activity, there are two main advantages for including analysis tools as part of a prediction activity. One is that the analysis tools hopefully help players engage directly with the domain data. The other advantage is that analysis tools may serve as a way to adjust the difficulty of the prediction activity by enabling and configuring which tools are available to the players during the activity. For example, the main prediction activity for Fantasy Climate used in the evaluations (chapter 6) would be more challenging without Thermovizz or the regression line feature of Thermovizz. That is because to make good predictions, the players would have to gather and analyze data on their own to discover long term trends in the weather data.

Another way prediction games support players in their prediction making is by including information resources that tie the prediction activity to the players' real-life context. During the prediction activity, the system queries an information source (e.g. a search engine, news aggregator, or social media platform) using keywords or phrases provided by the activity creator to retrieve information that may be valuable to the players during their prediction making task. As described in section 6.2, we explored alternative presentations of domain news in Fantasy Climate. As with the data analysis tools, tools providing domain content beyond the data used in the scoring formulas are also not required for the prediction activity. It is up to the activity author, whether to include them or not.

The configuration of the tools to analyze domain data and information resources for a prediction activity affect the user interface for the players. In Fantasy Climate, the "Tools" menu in the navigation bar and the "Data Analysis Toolkit" section in the prediction interface only show the tools that have been included and configured for the prediction activity (Figure.) These interface components become hidden if the activity author decided not to include them, as shown in Figure 29.

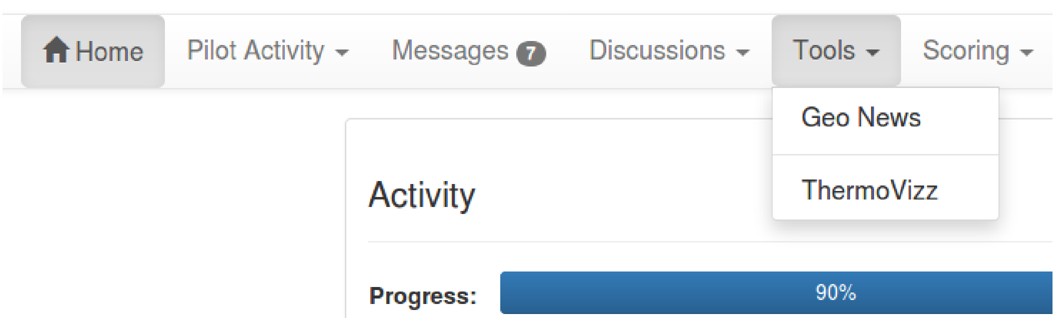

Figure 28. The 'Tools' menu and the 'Data Analysis Toolkit' showing the tools that have been configured for the prediction activity

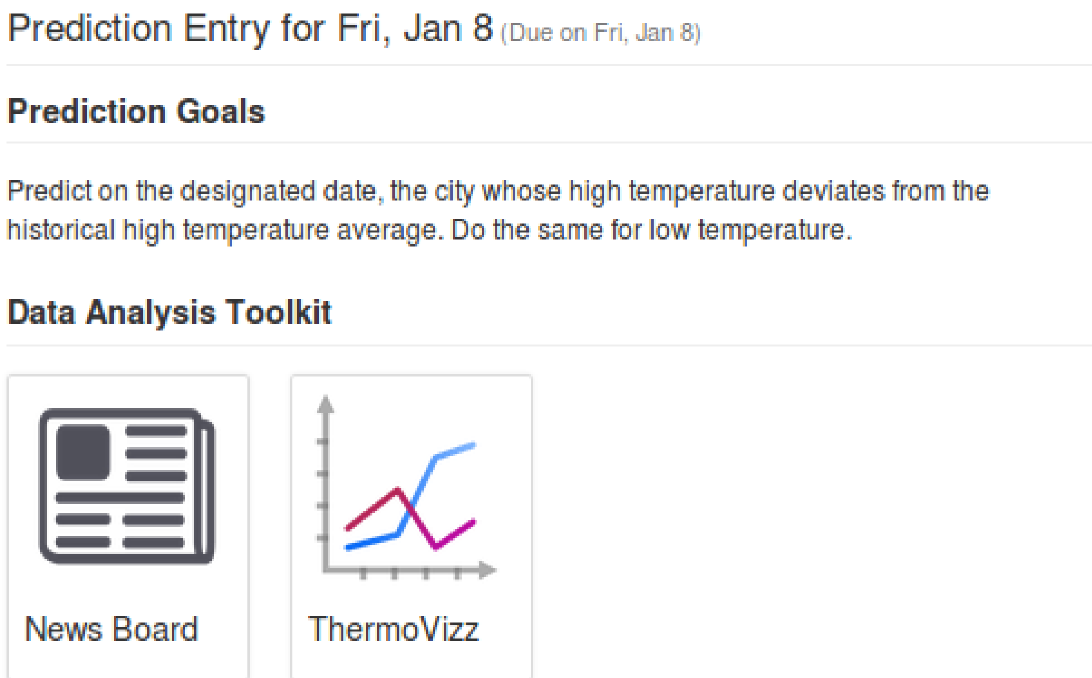

Figure 28. Continued.

8.6. Facilitating Social Interaction

The survey in section 3.1 and the evaluation in section 6.1 shows that social interaction plays an important role in fantasy sports and in prediction games respectively. To foster community building, our prediction game environment features a leaderboard for competition, and various ways for players to communicate with each other such as an instant messaging application, a comment wall, a topic-based messageboard and direct messages (see section 5.3). Like the analysis tools to support prediction making, these social interaction features are not required for the environment to run a prediction activity. The activity author decides and configures how the activity members communicate with each other and whether or not they compete with each other. For example, in a formal learning setting, an instructor may want to disable the competition feature to encourage more collaboration in her prediction activity. Figure 29 illustrates the main interface of Fantasy Climate for a prediction activity with no prediction support tools and all the social interaction features and competition disabled. All the social components (e.g. instant messaging) including their related menus (e.g. 'Discussions' menu) are absent. In such a setup, the “Scoring” menu that allows the players to access the full leaderboard is absent from the navigation bar, meaning players only see their own performance.

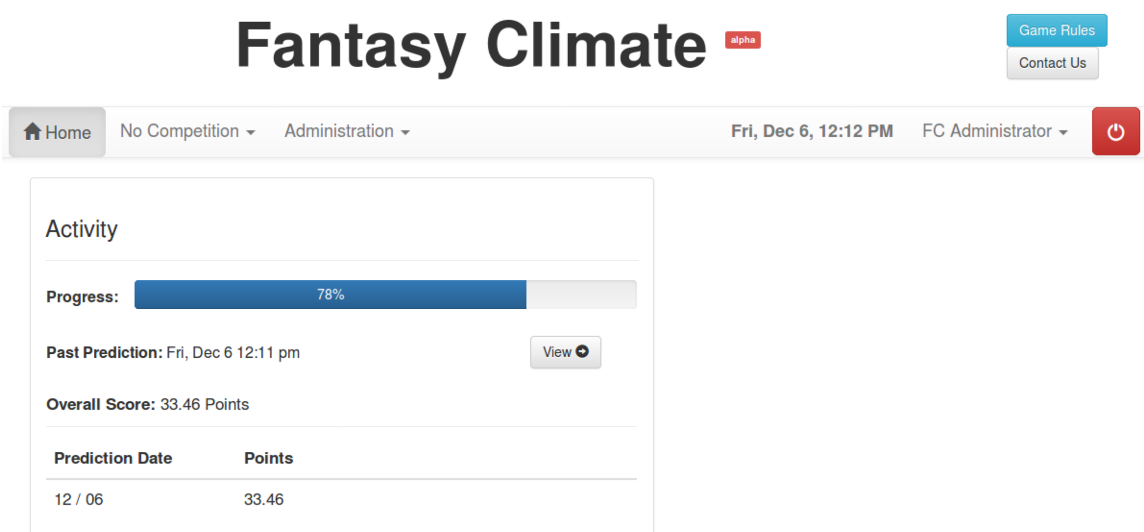

Figure 29. The main interface of Fantasy Climate for a prediction activity with no tools and, all social interaction features and competition disabled

8.7. Activity Members

The prediction game environment requires the activity author to specify who can participate in her custom prediction activity. In a casual setting, given the social nature of prediction games, the activity author may limit participation to friends or peers with similar interests. In a formal learning setting, the activity members of a prediction activity created by the instructor may be limited to the students in a particular course. Finally, one advantage of the control over the membership for a custom prediction activity is the ability to vary the support available for different players. For example, an instructor who wants some students to do more of the data analysis themselves may turn off the regression-line presentation of Thermovizz or Thermovizz overall for some students while maintaining those features for other players. In this way, the tools are a form of activity scaffolding that can be removed.

8.8. Authoring Framework Summary

The above discussion describes the minimum components and optional features that are part of a prediction activity definition. As they need to be customized for each distinct prediction activity, they inform the development of the environment for authoring prediction activities described in the next section.

9. AUTHORING ENVIRONMENT FOR PREDICTION ACTIVITIES: THE FANTASY CLIMATE CASE

The Activity Creation Wizard (ACW) is an environment that guides the author through a series of steps to author their prediction activity. Each step allows her to customize one or more of the prediction activity specifications described in the previous section. Hence each step may comprise more than one interface. The steps are:

1. *Defining the prediction activity objectives* – the author provides the activity name and the activity directives for the players
2. *Building the prediction activity schedule* – the author indicates when the activity starts, when it ends and, when the players' predictions shall be due and scored
3. *Creating the selection sets* – given a pivot set, the author constructs the selection sets from which the players will make their prediction in each the prediction round. In Fantasy Climate, the pivot set is the set of U.S. locations and a selection set may comprise a number of these locations (e.g. east coast coastal cities.)
4. *Defining the scoring rules* – the author creates or edits the formulas for scoring the players' predictions
5. *Selecting support for prediction making* – the author selects which support tools (e.g. Thermovizz and GeoNews in Fantasy Climate) will be available to the players during the activity and customizes the support they provide

6. *Customizing community interactions* – the author decides whether or not players compete with each other and how players communicate with each other

7. *Identifying activity members* – this is the final step; the author specifies who will take part in the prediction activity.

The design of the ACW aims to support activity authors throughout the authoring process. The ACW features a template catalog where an author may browse and search for prediction activity template that meets her needs. Also, at the end of the prediction activity authoring task, the author may save her customizations as a template which will then appear in the template catalog (Figure 32). In a formal learning setting, we imagine the template catalog to be a community repository where educators collaborate, seek or share the template of previously successful prediction activities authored by them or their peers.

Furthermore, the ACW provides explanations, rationales and examples for each authoring step and how the customizations therein will affect the prediction activity. The first time an author creates and customize an activity, the ACW encourages her to review the rationales and examples for each step by automatically popping up the explanations before the author can interact with the user interface, as shown in Figure 30. Later, the author may access explanations anytime during the task by clicking the help icon next to either a user interface section title, or a specific component therein (Figure 31.) The goals of this explanation system is to lessen (or even eliminate) the need for knowledge about prediction games, or more general computing knowledge when using the ACW.

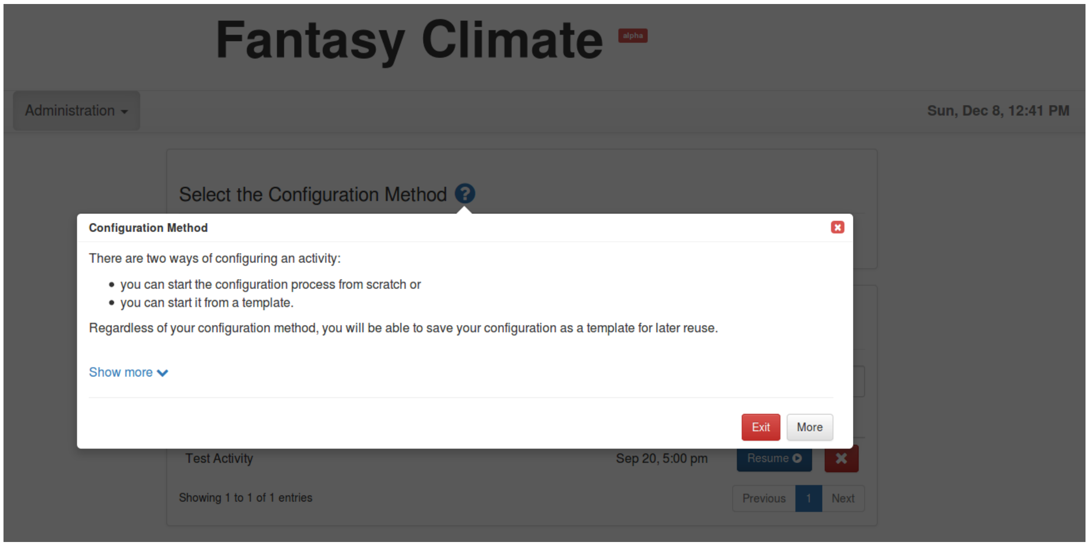

Figure 30. ACW popup explanations dialogs

Additionally, the ACW allows for a prediction activity authoring task to be interrupted at any time and resumed later. In every step, the customizations of the author are saved. When she later resumes the task, the ACW starts with the user interface of the step last accessed. Especially in a formal learning setting, we expect that it would be common practice to author a prediction activity in multiple stages. For example, before the semester, an educator may create and begin customizing a prediction activity that will be part of her course. But she cannot complete 'Step 7' of the ACW because she does not know who and how many students (future activity members) will be in the course hence she stops at 'Step 6' and saves her configuration. A week after the semester has begun and the student body of the course is a little bit more stable, she resumes the activity authoring task by completing 'Step 7' which is setting up the activity members.

Lastly, there are two steps of the ACW that we are anticipating may be tedious and repetitive during the activity authoring process: building the prediction schedule (step 2) and creating the selection sets (step 3). For example, building the schedule of a prediction activity that lasts a year and where predictions are made every two weeks could require setting 26 due dates and 26 scoring dates; a total of 52 entries. Thus, the ACW features a tool for building the prediction schedule automatically, and another tool for automatically generating the selection sets.

We will now describe the ACW support for initiating the authoring process and for each of the seven steps defined above.

9.1. Authoring a Prediction Activity

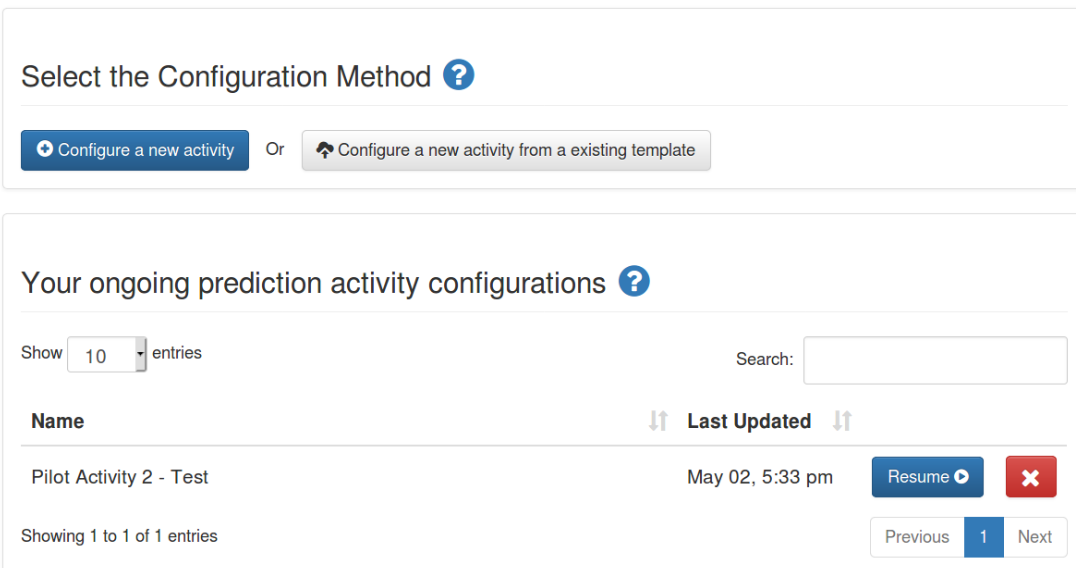

Figure 31. ACW configuration method interface

There are two ways of authoring a prediction activity. As Figure 31 shows, the author can start the authoring process from scratch by pressing the “Configure a new activity” button. This means that every customization of the prediction activity will be configured by her. Or she may press the “Configure a new activity from an existing template” button to start from an existing template. When customizing a prediction activity, the author can stop and save her progress at any step of the way. Authors may have multiple prediction activities; such as different activities for different classes or skill levels. The bottom portion of the interface in Figure 31 lists the user's ongoing prediction activity authoring tasks along with buttons to either resume or delete an ongoing task.

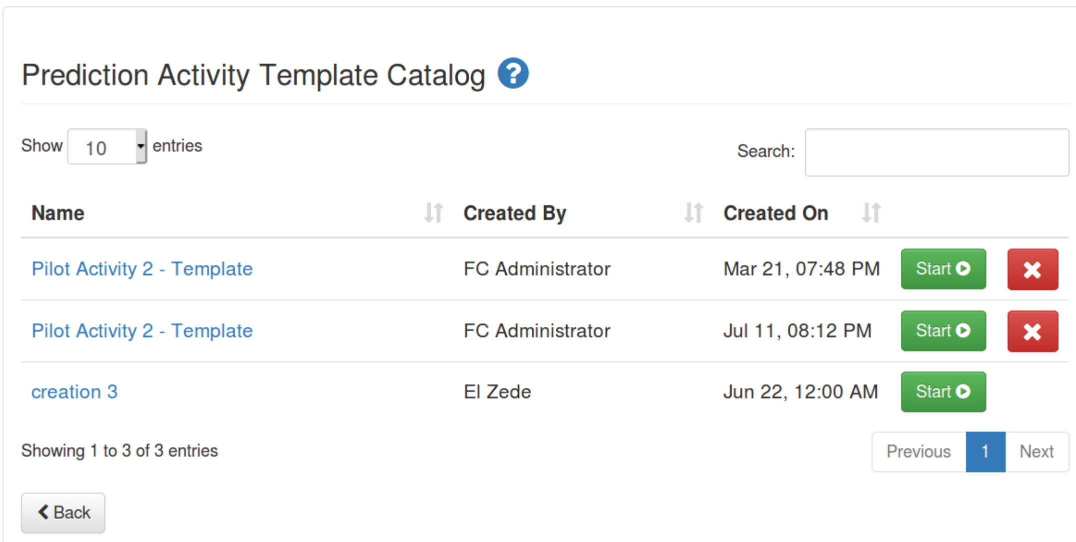

Figure 32. ACW template catalog

Once the author chooses to start the authoring process from a template, she is redirected to the template catalog displayed in Figure 32. The template catalog is the repository for existing activity templates. Clicking on the name of any template in the grid shows its prediction activity objectives along with the comments by its creator. The author can start the authoring task from a template by pressing the 'Start' button in the template’s row. She will still go through the entire authoring process. However, the interface(s) in every step of the ACW will be preloaded with the customizations of the selected template. Thus, she can focus on updating the components required to adapt the template to her activity objectives (e.g the prediction schedule). The catalog also allows the deletion a prediction activity template by its author.

9.2. Steps of the Activity Creation Wizard (ACW)

The Activity Creation Wizard (ACW) enables the authoring of a prediction activity one step at a time. There are features that are common to every step:

- the task progress – this interface component shows the completed steps, the current step, and the uncompleted steps,
- the navigation buttons – they are buttons that enable the author to move forward to the next step, backward to a previous step, or save and exit the ACW. The author may also use the task progress component to navigate to a specific completed step (e.g. the user is in step 6 and would like to edit step 2),
- validations – the ACW performs a validation whenever the prediction activity configuration is saved in the system (when any of the navigation buttons is pressed), displays the errors if there are any, and make suggestions on how to fix them. This ensures activity integrity and prevents identifiable errors (e.g. a prediction due date in the past or a misspelled function in the scoring formula) in the prediction activity configuration.

9.2.1. Step 1: Defining the Prediction Activity Objectives

Figure 33 shows the interface of 'Step 1' of the ACW that enables the author to enter the prediction activity name and its directives.

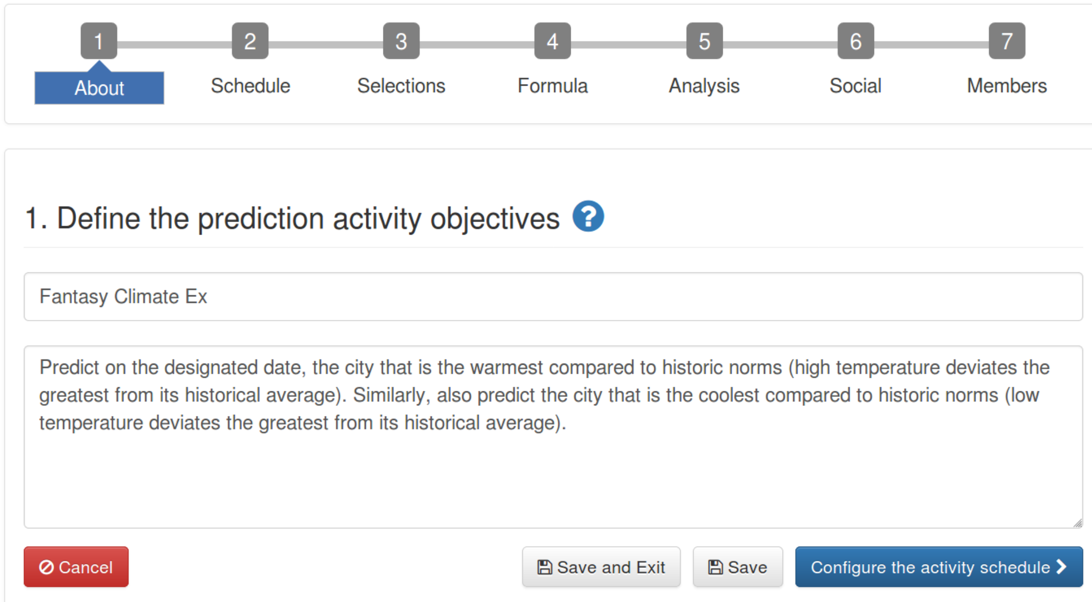

Figure 33. The ACW 'Step 1' interface

9.2.2. Step 2: Building the Prediction Activity Schedule

The interface in this step of the ACW lets the author set the start date, the end date, and configure the prediction rounds that each include the due date and scoring date of the prediction round's activity. Figure 34 illustrates the schedule of a prediction activity that will last about a month with two prediction rounds separated by a week.

Our prediction games system sends two types of notifications to players:

- prediction submission reminders – which are messages reminding them to submit their predictions before the prediction round due date and,
- scoring notifications – which alert the players that their predictions have been scored and that the results have been posted.

The author of a prediction activity may configure the prediction submission reminders in this step. Reminders are not required for the prediction activity to work. In such a situation, the players receive no notifications about the prediction round deadlines.

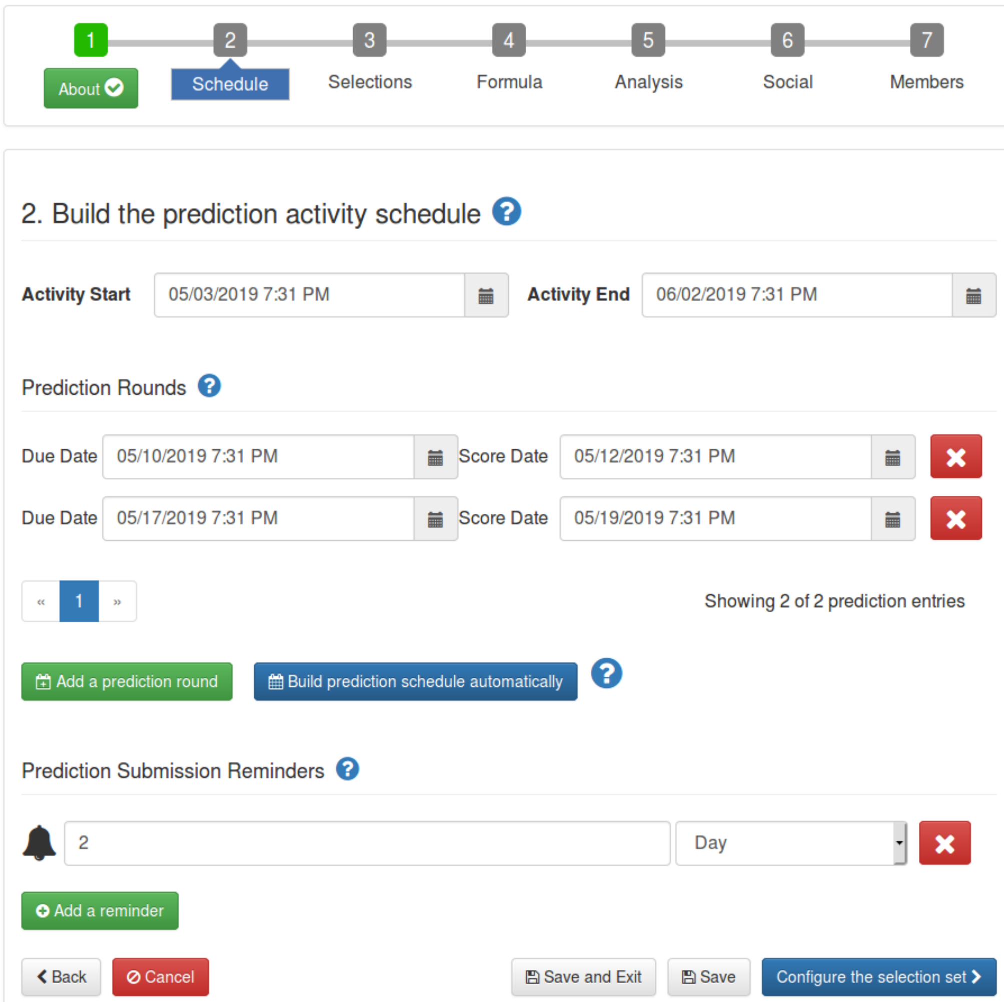

Figure 34. The ACW 'Step 2' interface

As mentioned earlier, generating a prediction schedule for a longer activity can be tedious and repetitive. The ACW provides to the author the option to automate the scheduling process using the automatic schedule builder shown in Figure 35. This tool

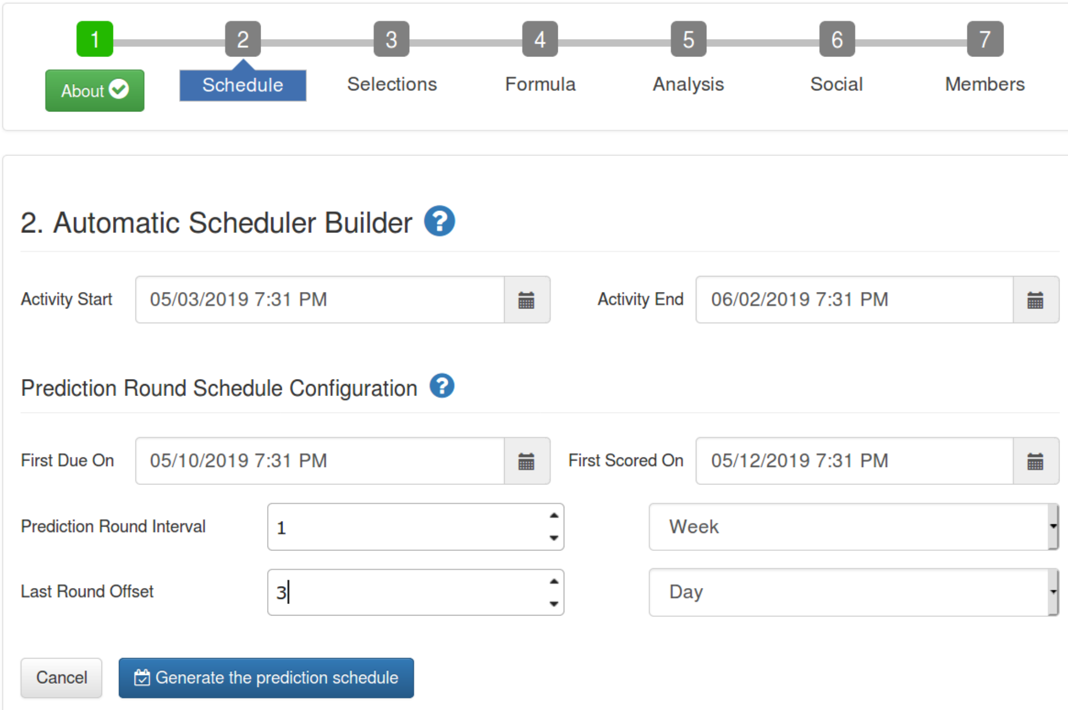

Figure 35. The automatic schedule builder interface

reduces the efforts of setting up a prediction by automatically generating a schedule based on the following inputs:

- the prediction activity schedule boundary: the start date and the end date
- the first prediction round: the first due date and the first scoring date
- the prediction round interval: this is the length of time between prediction rounds (e.g. 1 week). In other words, it is the frequency of prediction rounds in the activity schedule. For example, to set up an activity with weekly prediction deadlines (due dates), the prediction round interval value must be set to 1 week
- the last round offset: this is the length of time between the last scoring date and the activity end date (e.g. 3 days in Figure 35). This period is meant to give

players time to review and reflect on their performance (and that of their peers) and communicate with other participants before the prediction activity officially expires.

The generated schedule will appear on this step's main interface (Figure 34) for further editing.

9.2.3. Step 3: Creating the Selection Sets

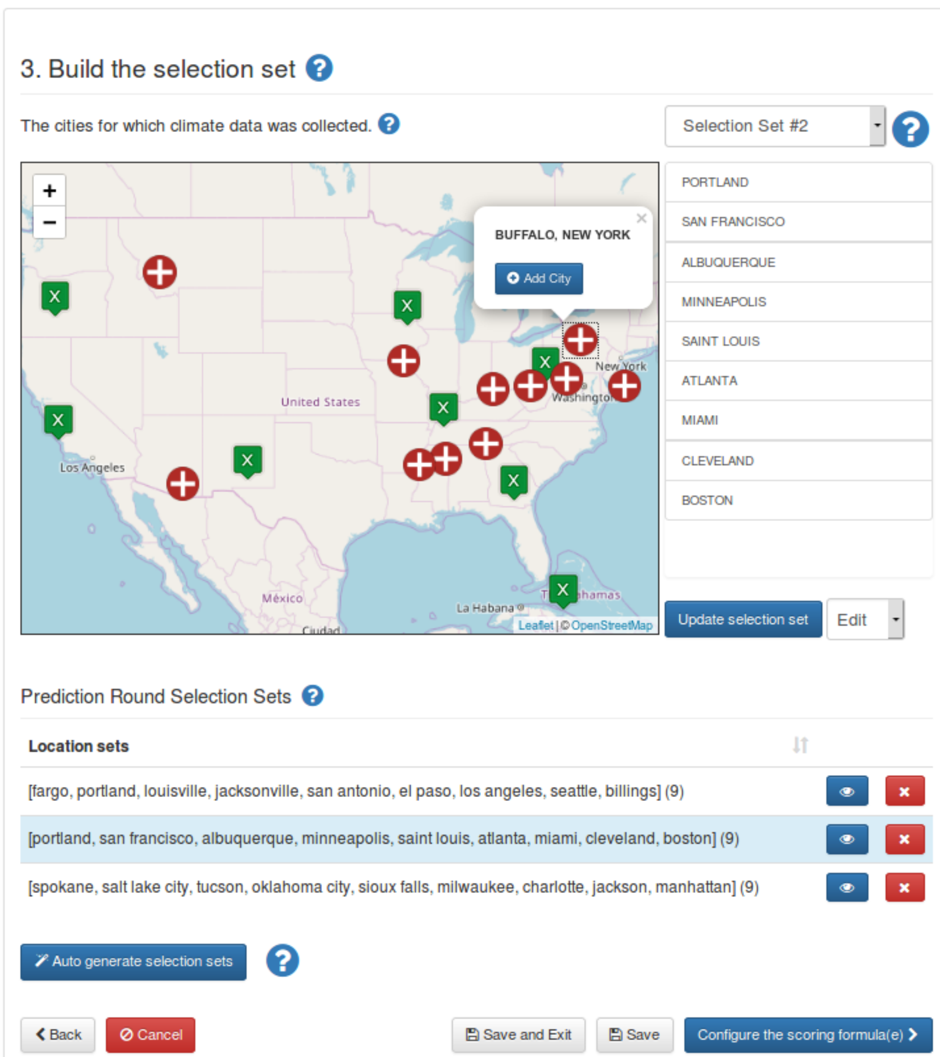

Figure 36. The ACW 'Step 3' interface

Predictions in activities can take different forms. In Fantasy Forecaster, described in section 3.2, predictions were quantitative values (e.g. high temperature) and multiple choice selections (e.g. “partly cloudy”.) In Fantasy Climate, predictions involve participants selecting among varying sets of locations based on predictions about how upcoming observed temperatures will vary relative to historic norms. The definition of the selection sets is supported by the interface in Figure 36. This interface is appropriate for Fantasy Climate or any other prediction games in domains where the pivot set is made up of U.S locations. The map component visualizes the pivot set with markers that enable the author to add/remove a location to a new selection set or a highlighted one. In Figure 36, the author is in the process of adding the city of Buffalo, New York to the second selection set. The table below the map component shows the 3 selection sets created for the prediction activity; each set containing 9 cities.

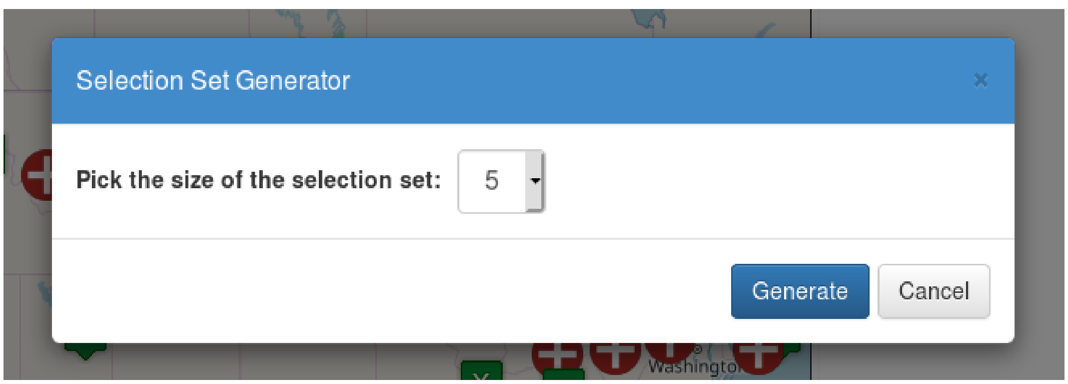

Figure 37. The selection set generator interface

Pressing the “Auto generate selection sets” button at the bottom of the interface launches the selection set generator displayed in Figure 37. This tool automatically generates several selection sets of the size picked by the activity author (e.g. the size of 5 in Figure 37.) The generated results will appear on this step's main interface for further editing.

9.2.4. Step 4: Defining the Scoring Rules

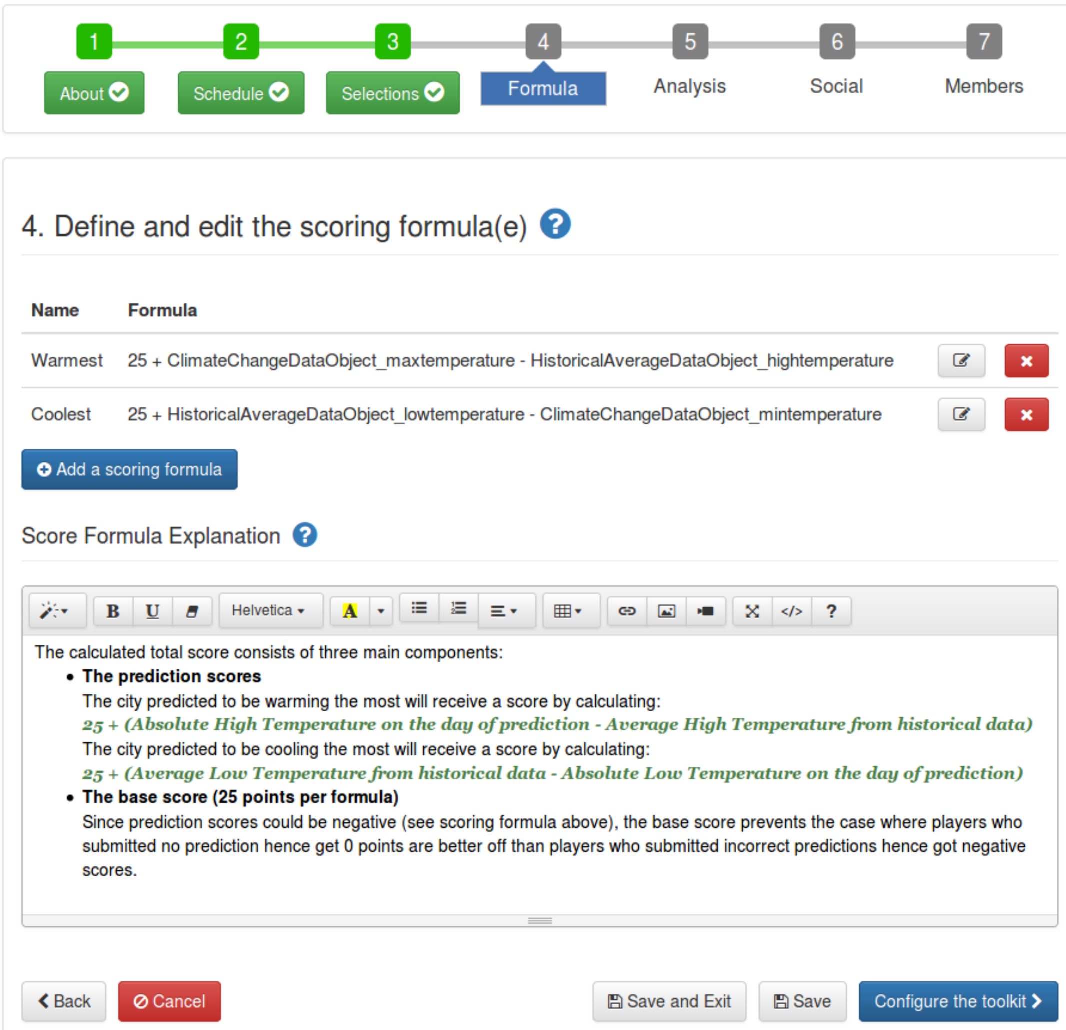

Figure 38. The ACW 'Step 4' – the formula viewer

The formulas used to score players' predictions are created in this step. It has three interfaces: (1) the formula viewer, (2) the data property picker, and (3) the formula editor. The formula viewer in Figure 38 contains a table that lists the prediction activity scoring formulas. The first column represents the scoring formula's prediction name, and the second column is the actual scoring formula. Variable names in the formula are formatted as {dataset}_{property} where 'dataset' is a domain dataset (e.g. ClimateDataObject or HistoricalAverageDataObject) and, 'property' is a data property of this dataset (e.g. maxtemperature.) This method of identifying elements will be understood by those familiar with data structures in most programming environments but may require support for those without such experience.

The formula viewer also encourages the prediction activity author to explain their scoring formula. Score explanations like the one in Figure 38 are meant to provide transparency, clarifying to the players how their submitted predictions will be evaluated. During the prediction activity, the explanation content will appear in a dialog box launched by clicking on the help button (e.g. question mark) adjacent to the player's score. To navigate to the formula editor, the prediction activity author may press either the “Edit” button next to a formula or the “Add a scoring formula” button.

Before editing the formula, the activity author selects features available to be included in the formula via the domain data property picker, shown in Figure 39. Some datasets may have many properties (e.g. weather data may easily have more than 15 features per location per observation). The domain data property picker helps the activity author focus on the aspects (e.g. temperature) of the dataset (e.g. weather) that are central to her scoring formula and subsequently to the prediction activity. The interface is a view of the prediction game domain datasets metadata: the dataset name and the dataset properties. In the current implementation of Fantasy Climate, there are two domain datasets:

- the weather dataset named *ClimateChangeDataObject* with properties: high temperature, precipitation, wind, etc..., and
- the historical average weather dataset named *HistoricalAverageDataObject* with properties high temperature, precipitation, wind, etc... which are the averages of weather properties (see chapter 6.)

In Figure 39, the author has picked the property *maxtemperature* from the *ClimateChangeDataObject* dataset and, the property *hightemperature* from the *HistoricalAverageDataObject* to be used in her scoring formula.

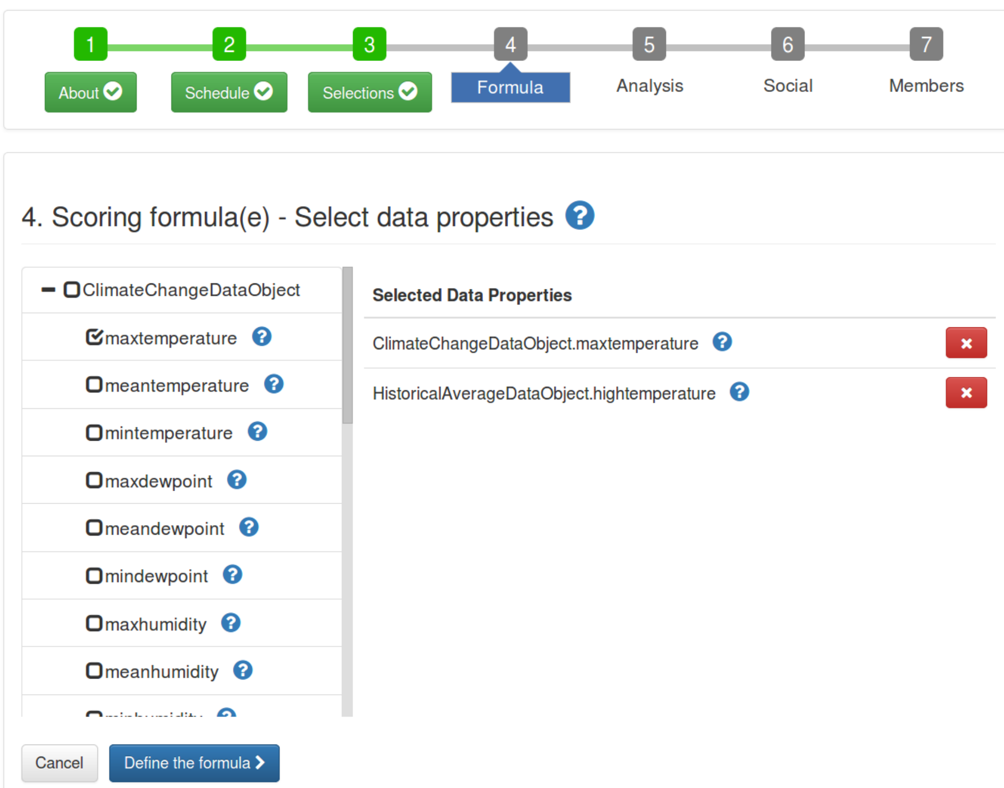

Figure 39. The ACW 'Step 4' – the data property picker interface

Once data properties have been picked, the author navigates to the formula editor (Figure 40). In this interface, she provides the prediction name and the mathematical expression of her scoring formula. She can access her previous picked data properties under “Data Properties” menu, and various mathematical functions (e.g. square root, absolute value, etc...) under the 'Math Functions' menu to incorporate during the expression of her formula.

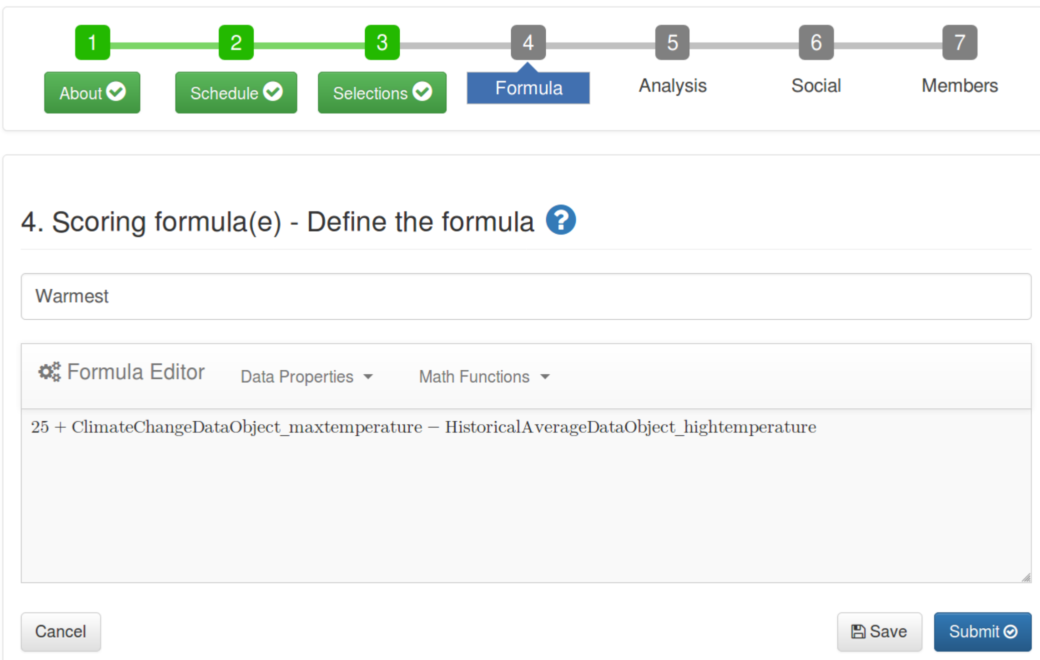

Figure 40. The ACW 'Step 4' – the formula editor interface

9.2.5. Step 5: Supporting Prediction Making

In this step of the ACW, the author configures the kind of support that will be available to players for their prediction making during the activity. This is a two-step process involving the selection of data analysis tools and information resource tools.

Figure 41 depicts the data analysis tool catalog that lists all the tools that are either applicable to, or have been specifically developed for, Fantasy Climate. For example, Thermovizz was developed specifically for Fantasy Climate. Whereas the Data Exporter, which allows players to download data from the activity for use in other applications (e.g. Excel), can apply to other prediction games in other domains. Clicking on the tool icon shows more details, including the options to enable and configure it. In Figure 41, Thermovizz has been highlighted and enabled for the prediction activity.

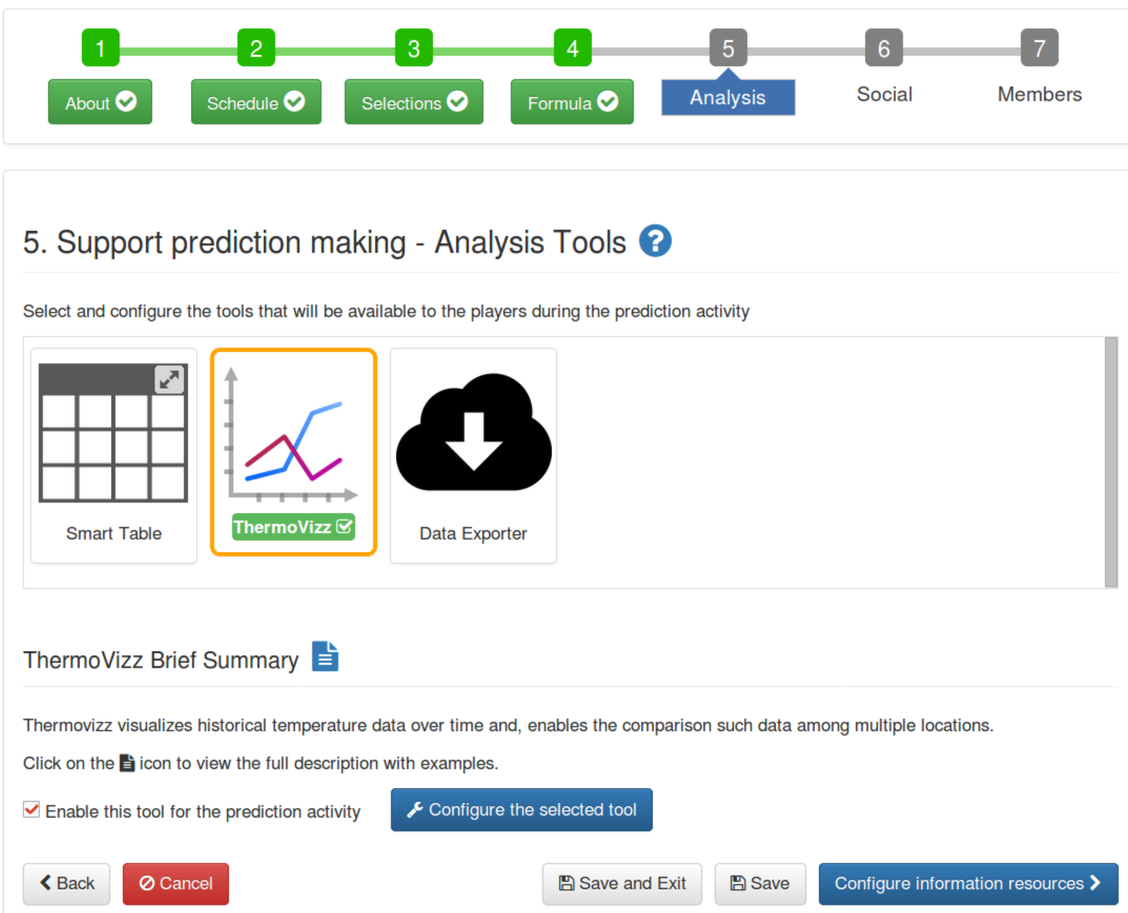

Figure 41. The ACW 'Step 5' – data analysis tools

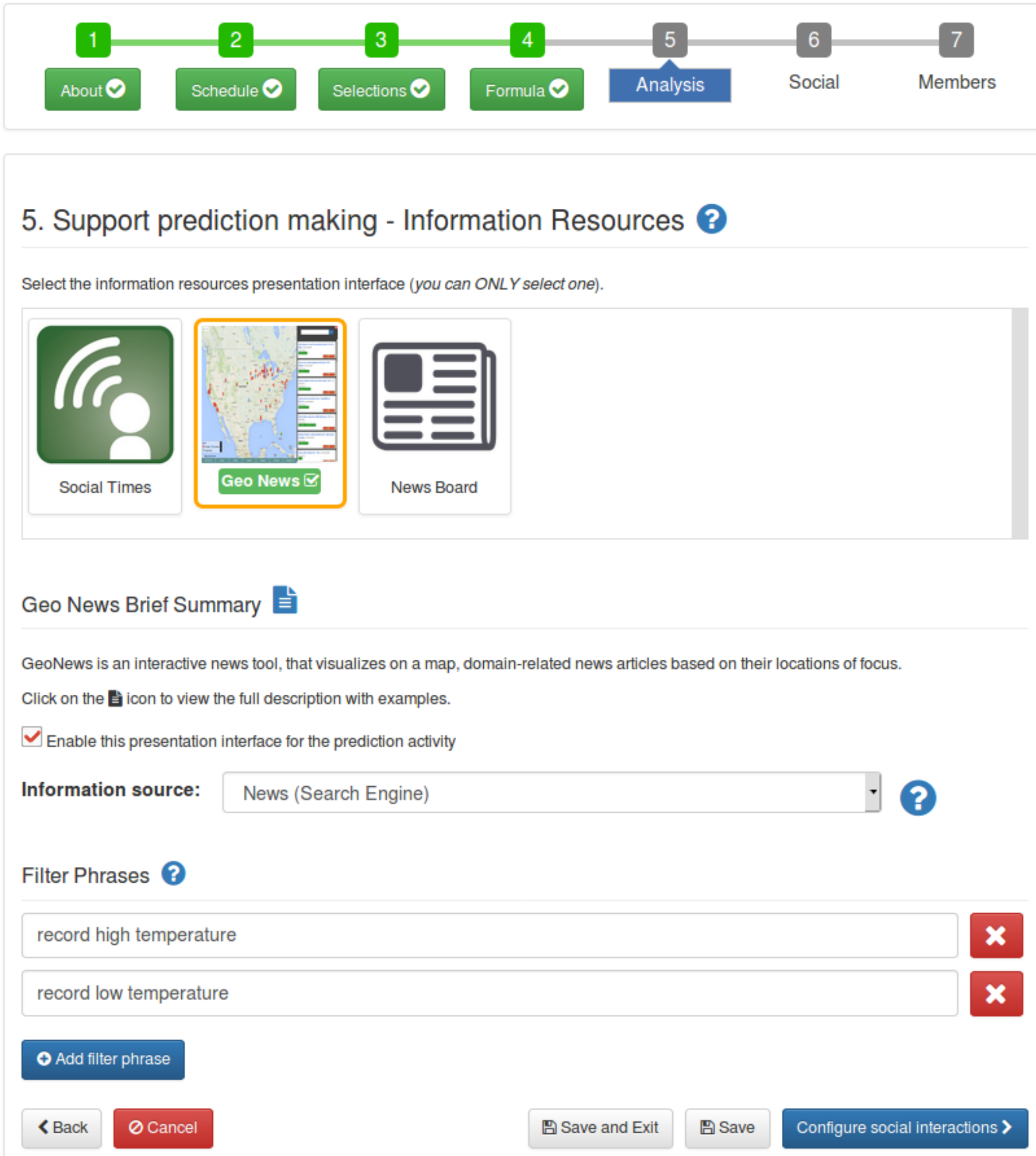

Figure 42. The ACW 'Step 5' – information resources

Figure 42 illustrates how the activity author configures the information resources tools (e.g. domain-related news) for the prediction activity. She begins by selecting and enabling the information presentation tool, in this case GeoNews (see section 5.2), from a catalog that functions identically like the data analysis tool catalog in the prior step.

Then she picks the source, in this case a search engine (e.g. news.google.com), from which the information resources will be retrieved. An alternative information source may be Twitter or Facebook. Last, the author provides the key phrases and keywords that the prediction game environment will be use to locate information resources that are likely relevant to the prediction activity. In this case the author is searching for articles that mention either "record high temperature" or "record low temperature."

9.2.6. Step 6: Customizing Community Interactions

Here the author decides which communication tools will be available, and whether players will compete with each other during the prediction activity. The four player communication mechanisms described in section 5.3, namely instant messaging (chat), direct messaging (in-game email), a topically-structured forum, and a temporally-structured comment wall, are the available options. In Figure 43, she has turned everything on for her custom prediction activity.

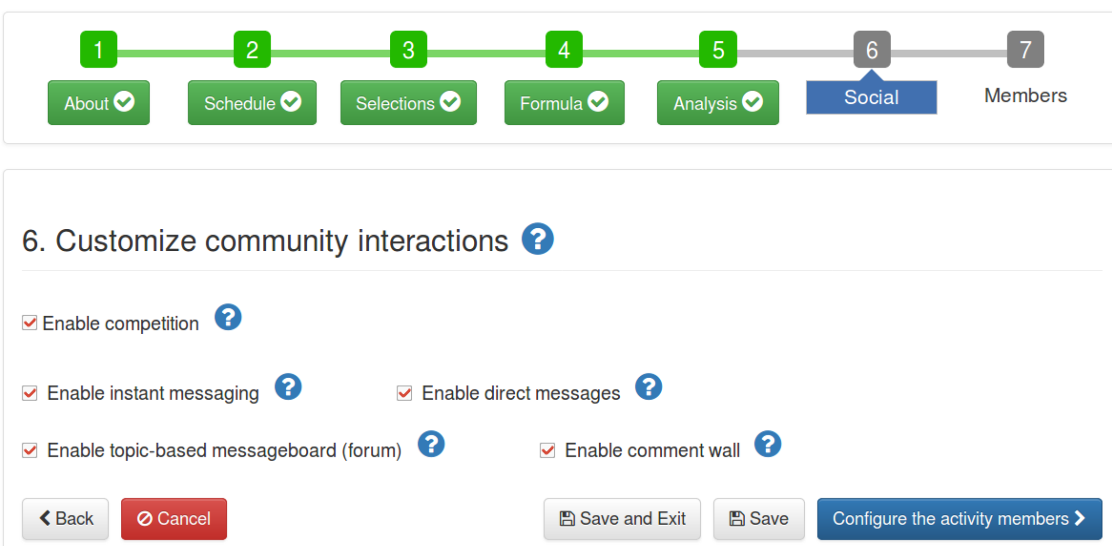

Figure 43. The ACW 'Step 6' interface

9.2.7. Step 7: Setting up the Activity Members

In the final step of the prediction activity configuration, the activity author selects a subset of known users that will be the players (activity members) in her prediction activity (Figure 44). If her desired players are not yet users in the system, she may navigate to the “Upload New Users” interface (Figure 45) to create them.

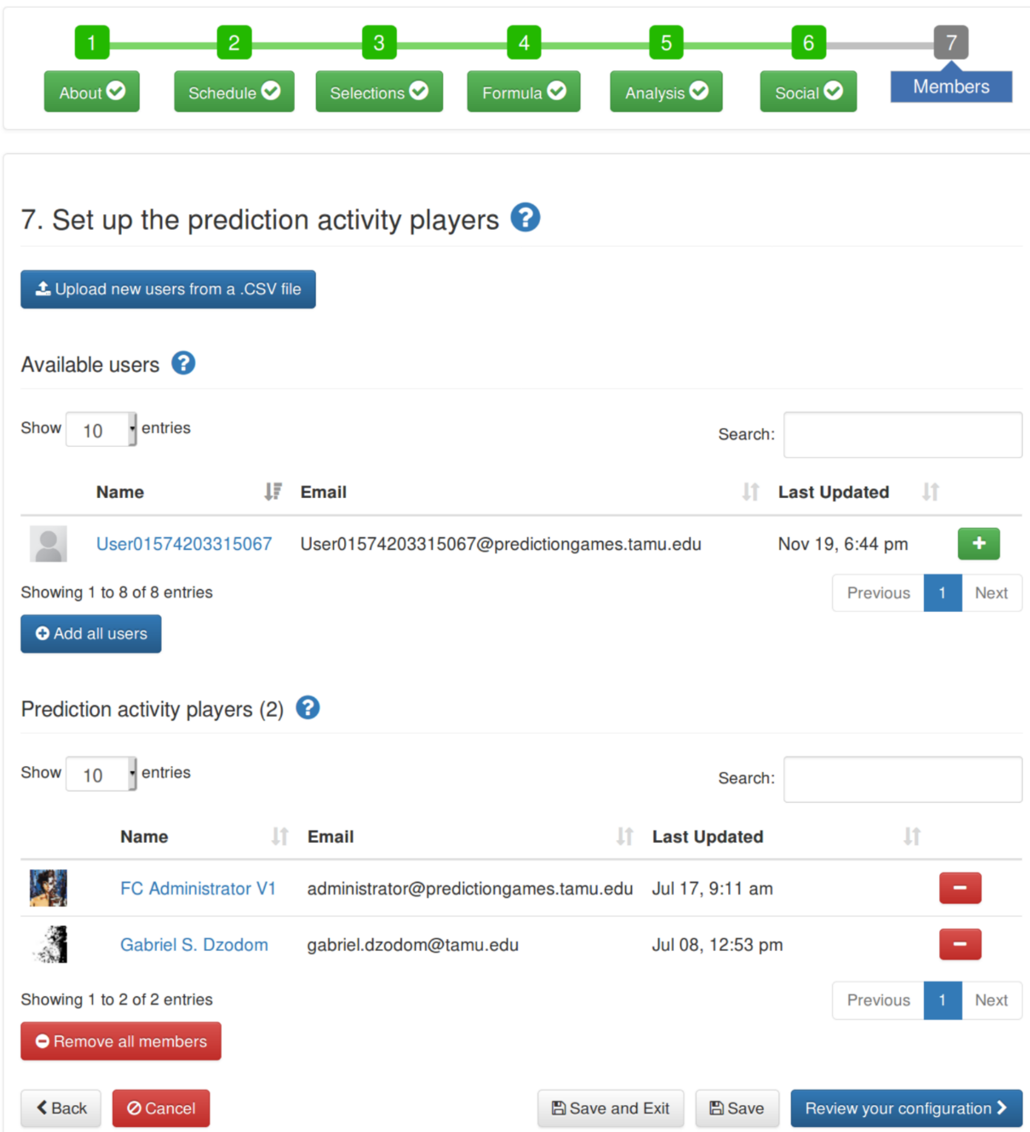

Figure 44. The ACW 'Step 7' interface

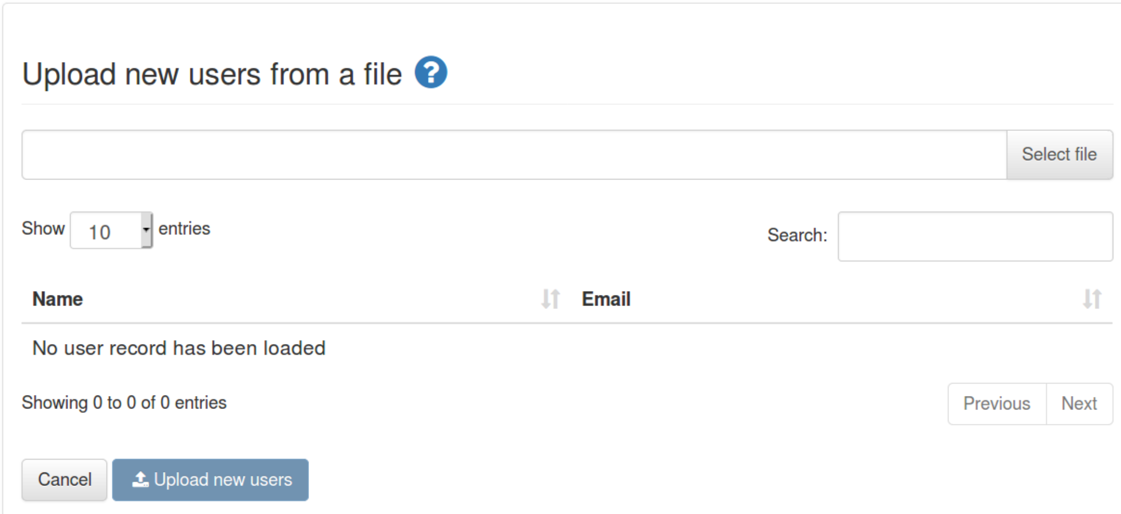

Figure 45. The ACW 'Step 7' - uploading new users from a file

9.2.8. Reviewing the Prediction Activity Customizations

Before the activity author generates a new prediction activity, the ACW presents a summary of the selections and customizations so she can review her work. Figure 46 shows the customizations for every step of the ACW. Each section of the summary corresponds to a step in the authoring process. Clicking on the 'edit' icon next to the section's title redirects to the corresponding step's interface for editing.

Once the activity author is satisfied with her customizations, she may first save it as a template which will be accessible from the activity catalog. Or she may submit her customizations directly. Once submitted, a prediction activity will be generated. It will be in a pending state until the start date is reached, after which the activity will truly begin.

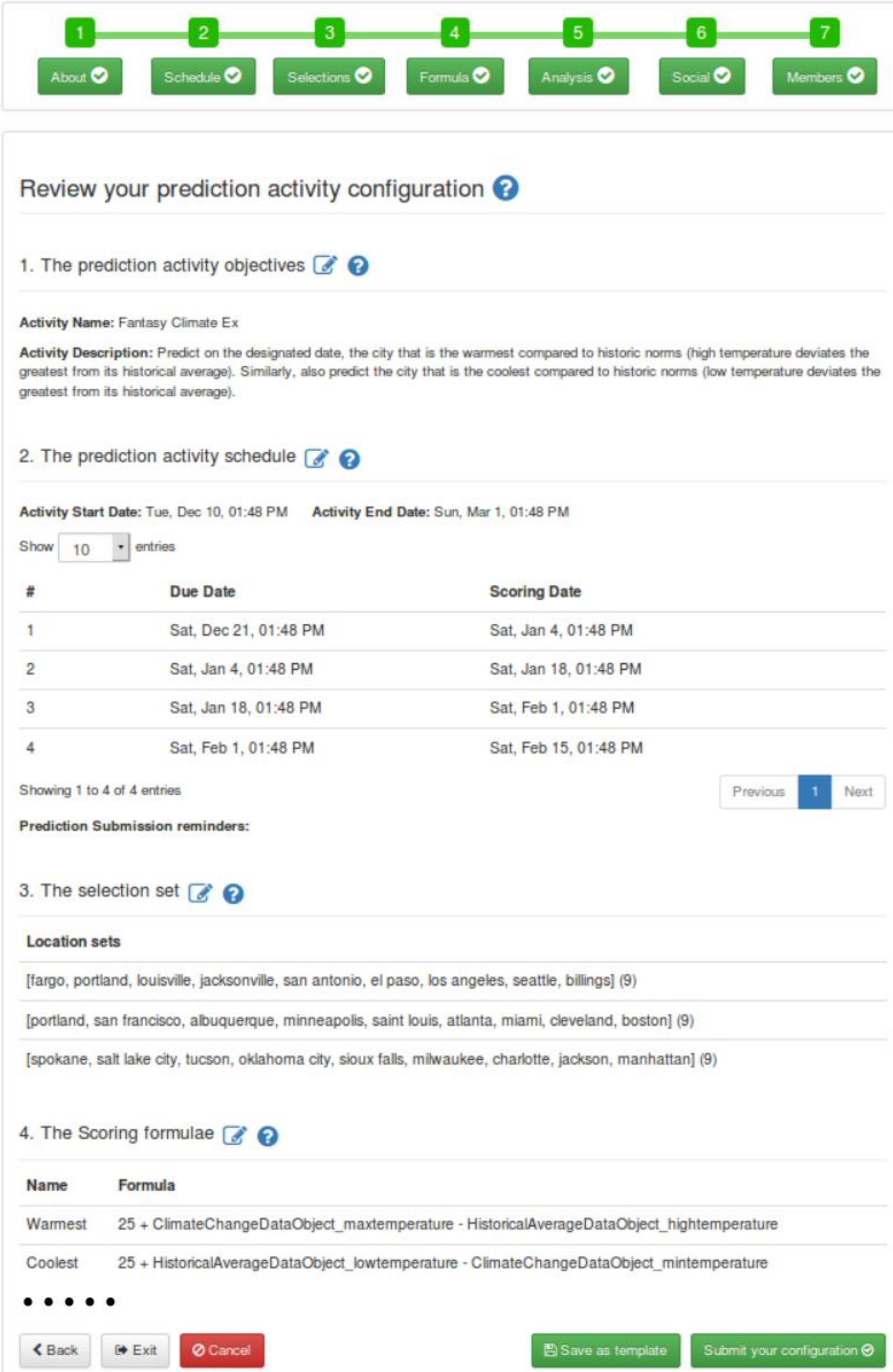

Figure 46. The ACW Reviewing interface

9.3. Supporting Fantasy Precipitation using the ACW

The specifications of a prediction activity as described in chapter 8 are meant to be minimal. Many prediction games will have specifications that are particular to their particular domain or type of prediction. As a consequence, the ACW steps and components may need to vary depending on a game's goals. In terms of features that are applicable to all prediction games, the name and directives (Step 1), the prediction schedule (Step 2), supporting social interaction (Step 6), and activity members (Step 7) are the more general in character. Whereas the selection sets (Step 3), the scoring formula (Step 4), tools supporting prediction making (Step 5), and the content of the explanations may be domain and game dependent. To give a sense of how this might require programming to revise/extend the ACW, we describe the necessary enhancements for Fantasy Precipitation (described in Appendix C.)

In Fantasy Precipitation, players predict whether precipitation will occur or not within a date range (e.g. 01/15 – 01/12) for every city of a given set. Fantasy Precipitation and Fantasy Climate are both based on weather and share the same dataset so historical and real-time data ingestion is not an issue. As a consequence, the current version of the ACW is capable of producing a Fantasy Precipitation activity with three exceptions that require some programming. The first exception is the scoring formula step because the players' predictions values and the values to evaluate them do not come from the domain data properties (e.g. high temperature, precipitation) as they do in Fantasy Climate. In Fantasy Precipitation, the predictions' values are Boolean selections (True or False for whether there will be precipitation in a location) and the values to evaluate them are generated by code based on the domain data properties provided by the external data source. Thus, some programming mapping all non-zero values for rainfall and snowfall would be needed given the current state of the ACW. Figure 47 shows the data selection interface of the scoring formula step for Fantasy Precipitation.

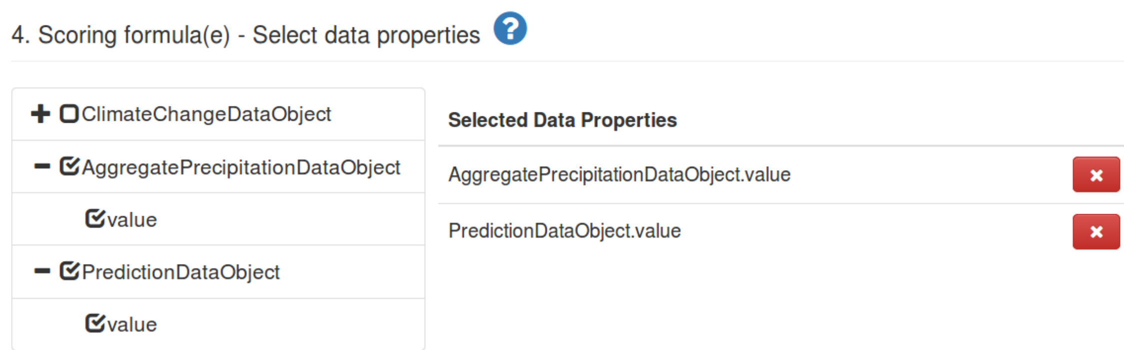

Figure 47. The data selection interface of the scoring formula step for Fantasy Precipitation. 'PredictionDataObject' represents the user prediction's data and AggregatePrecipitationDataObject represents the data used to evaluate the user predictions.

The second exception is the addition of a tool into the tool catalog (Figure 41) aimed to more specifically support predicting precipitation for a location given a date range. The ACW's current version does not feature a tool for visualizing precipitation data. Such a tool would have to be programmed and incorporated in the catalog. Figure 48 shows the tool catalog interface for Fantasy Precipitation including the Location Profile tool (Appendix C) that was developed for it. The ACW was developed to have an extensible list of tools to select among so once the tool was developed, it could be added to those available for selection by activity authors with minimal effort.

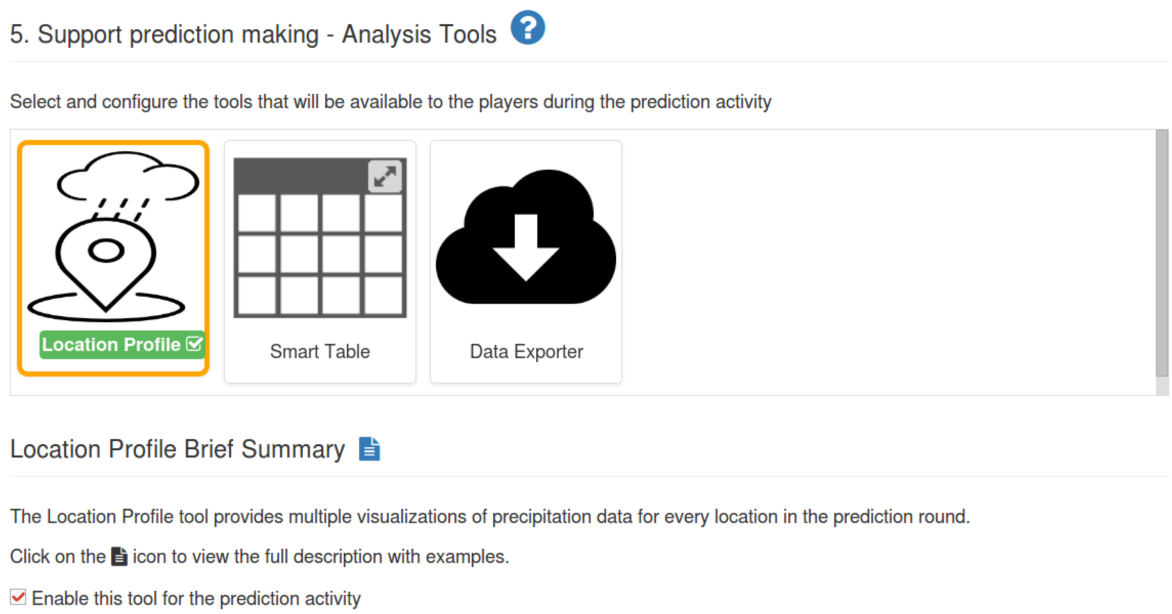

Figure 48. The tool catalog in Step 4 of the ACW for Fantasy Precipitation

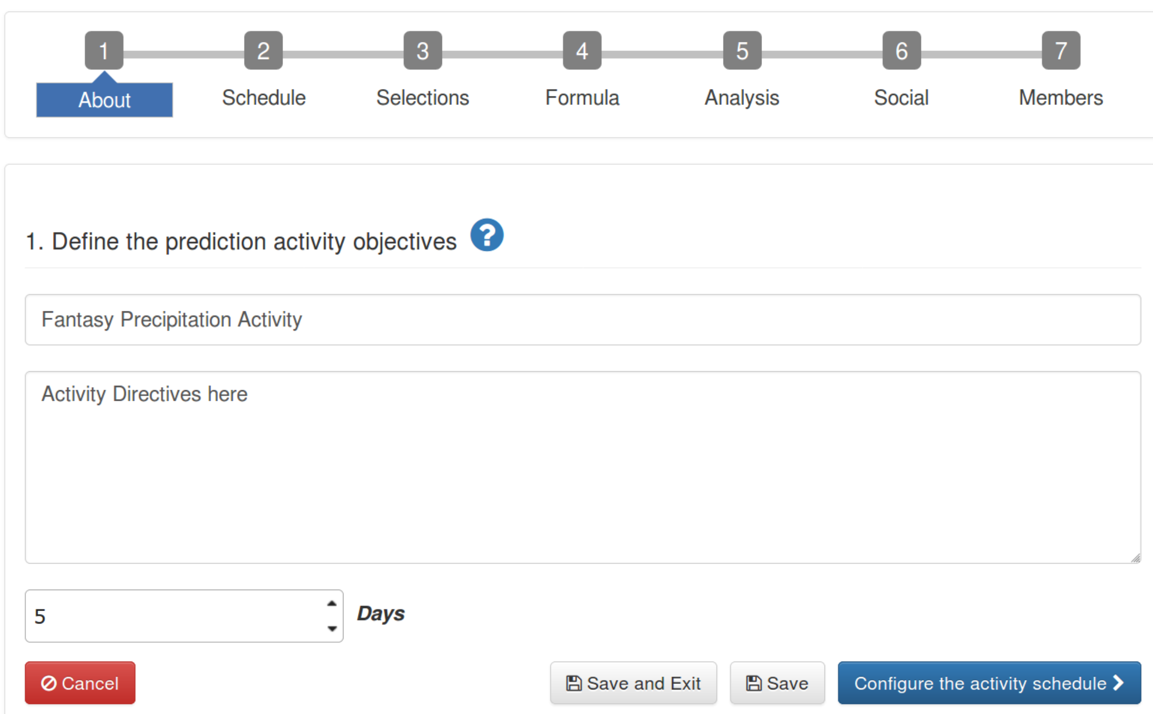

Figure 49. ACW Step 1 interface for Fantasy Precipitation

The final feature of Fantasy Precipitation that would require programming is enabling the activity author to set the length of the prediction date range (e.g. 5 days), which is a specification that is particular to Fantasy Precipitation. Figure 49 shows how it was implemented.

Thus, to adapt the ACW to Fantasy Precipitation does require programming for the three extensions described but many of the core prediction game components could be used without change. Additionally, once these additions were made to the ACW, these new capabilities could be used for yet other prediction domains and activities.

9.4. Authoring and Activity Creation Wizard Summary

In summary, the steps of the ACW described in the sections above include the minimum steps required to author a prediction activity but adds the ability to select among data analysis tools, information resources and communication tools for the activity. These general steps are applicable to most prediction game although the interfaces needed may vary depending on the domain and types of predictions desired.

10. STUDY OF PREDICTION GAME AUTHORING

The activity creation wizard (ACW) is an authoring environment that guides the user step by step in her task of creating and customizing a prediction activity. The ACW includes multiple components that facilitate the prediction activity authoring task. To help authors with the stepwise process is the explanation system that provides the activity author with context, explanations (including rationales) and examples. The explanations are expected to be particularly valuable for new activity authors. Another component is the template system that lets the author create a prediction activity based on an existing activity. A template is also an alternative way to provide examples of a prediction customization to the author. Furthermore, the ACW features a couple of components to automate repetitive tasks during the prediction activity authoring process: the automatic schedule builder to automate the prediction schedule and the selection set generator to automate to creation of selection sets. Thus, the primary goals of the ACW evaluation are:

- to understand how the different components affect the activity authoring process, especially the template system and the automation components,
- to gather feedback on the ACW (e.g. limitations, which steps are difficult to understand or operationalize, suggested improvements), and
- to better understand the authoring process of a range of authors (e.g. the order of the ACW steps, alternative ways of customizing the prediction activity).

The design of a prediction activity is a creative process. Deciding on a domain, determining what is an aspect of that domain to focus on with the predictions, and how to score the predictions all require consideration of the potential players, their expected knowledge, and goals for the activity. Additional goals of the evaluation are:

- to discover examples of prediction games people think would be interesting, and
- to better understand how authors conceptualize this high-level design task.

10.1. Experiment Setup

We conducted a study to explore the authoring of prediction activities and to evaluate the ACW in support of this process. The research questions were:

- How do different components of ACW affect the prediction activity authoring process?
- How does background experience affect the authoring process?
    - experience with fantasy sports
    - experience with online games in general
    - experience with statistics and data analysis
- What steps are the most difficult for the author?
    - Which steps take the most time (and why) during the authoring process?
- Is the steps order important for the authoring process?
- What are the limitations of the ACW?
    - What aspects of Fantasy Climate the user would like to customize but are not supported by the ACW?
- What types of predictions games users would like to create?

To investigate the effects of the template system and the automated support components on the activity authoring process, we setup for four versions of the ACW where these components were either included or not. These four versions map to the four conditions in the study. Table 2 describes these four versions and provides the codes that will be used in the results section when referencing to the system versions.

Table 2. The four versions of the ACW setup for the study

	Components	**Code**
Version 1	Template & automated components	T-A
Version 2	No template. Only automated components	NT-A
Version 3	No automated components. Template only	T-NA
Version 4	No template & no automated components	NT-NA

Participants in the study were recruited via email (Appendix E) and word of mouth from the university community (mostly students). The resulting 24 participants were evenly and randomly distributed to the versions of the ACW so that each version had 6 participants. None of the participants had any prior knowledge of the prediction games project. Each participant's prediction authoring activity occurred in the same office with the same computing infrastructure. The participant mainly completed two tasks: (1) to use the ACW to author a prediction activity for Fantasy Climate given relatively well-defined requirements, and (2) to design a prediction game in a domain of their choosing given limited requirements. The requirements for the tasks were constant for all the participants. Appendix F and Appendix G describes both tasks respectively.

We collected data through screen captures (task 1), questionnaires, semi-structured interviews, and content produced by the participant.

Once each participant arrived, settled in our office and signed the consent form (Appendix D), we began the study with an introduction to the prediction games project, and a demonstration of Fantasy Climate showing the components of the game and their functionalities. Participants were then offered the opportunity to interact with the game. Next, participants answered a short questionnaire (Appendix H) about their background (e.g. demographics, experience with fantasy sports, etc...). Participants then used the ACW to author a prediction game for Fantasy Climate. Upon completion of the task, the participant answered another questionnaire (Appendix I) followed by a short interview (Appendix J) to collect quantitative and qualitative information about her experience with the ACW. Afterward, we asked the participant to design their own prediction game using the medium of their choice (e.g. Word Processor, paper). Finally, we ended with a short conversation (Appendix K) about the participant's game and elicited further background information spurred by the design of their fantasy prediction game.

10.2. Demographics and Background

Before doing the first task, the participants answered a short questionnaire asking about demographics (3 questions), educational background (3 questions), and gaming background including their experience with fantasy sports (3 questions). The questionnaire combined single-answer questions, multiple-choices questions, and Likert-scale questions. Appendix H shows the details of the questionnaire.

The average of age of participants was about 26 years old (standard deviation = 4.64). The maximum age was 35 and the minimum was 18 years old. There was an equal number of female and male participants. Asian/Pacific Islander were the biggest racial group (16/24), followed by Caucasians (4/24). Black/African American, Hispanic/LatinX and, Multiracial has one participant each. And one participant chose not provide his/her racial background (Figure 50).

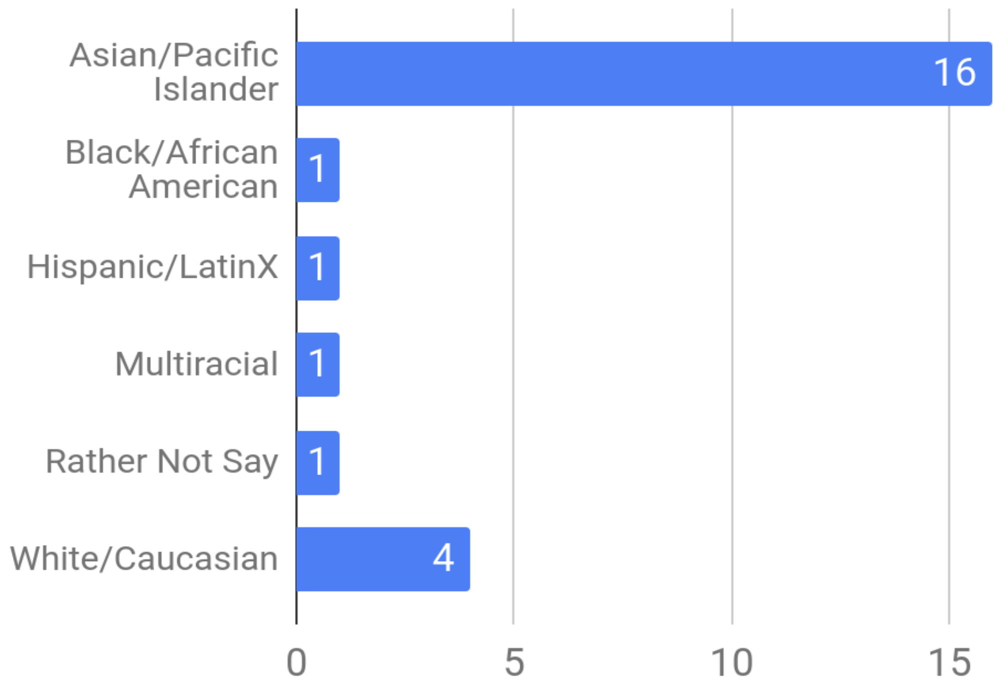

Figure 50. Racial distribution of the participants

All the participants were recruited from the university community. As a consequence, all the participants were students or postdoctoral researchers. 75% of them (18/24) were graduate students/postdoctoral researchers or had at least a bachelor's degree whereas 25% were undergraduates (Figure 51).

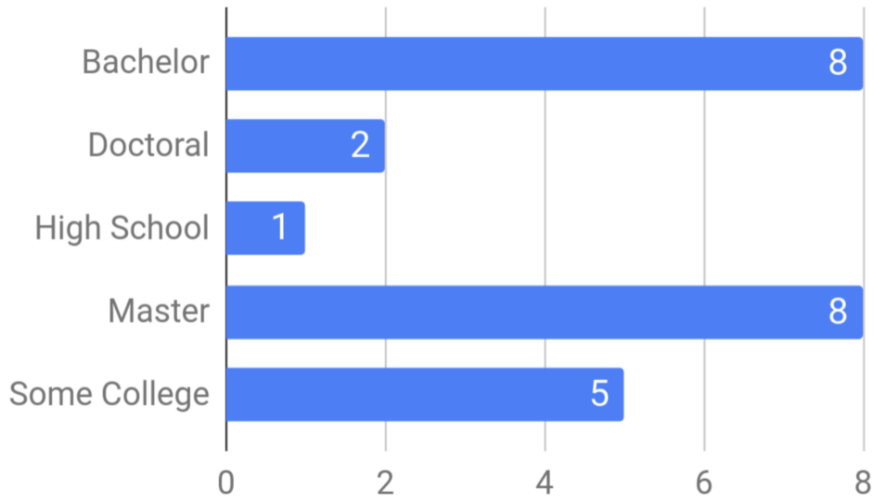

Figure 51. Highest level of education completed

In Figure 52, half of the participants had an engineering background and 5 had a science background. There were 2 participants with background in humanities and another 2 with background in business. The 2 'Other' responses were information technology and health science. The remaining participant had a background in architecture design.

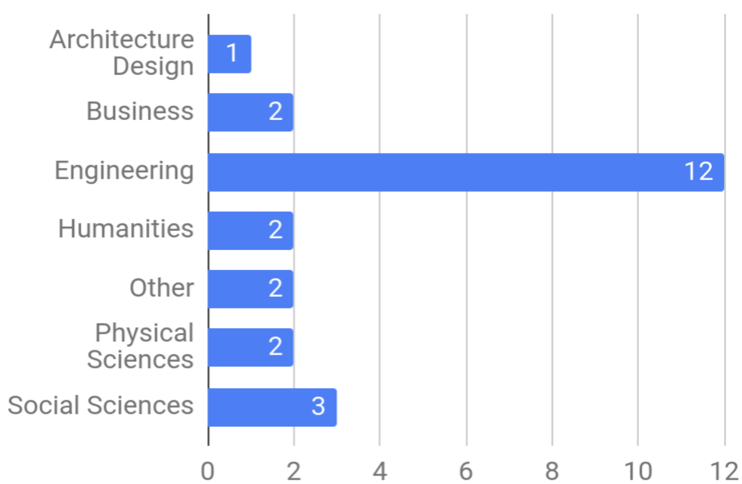

Figure 52. Fields of study background

When asked to rate their knowledge of statistics, no one reported being an expert and only 1 participant reported having no knowledge of statistics. The rest of the participants (23/24) answered having at least a basic knowledge of statistics as shown in Figure 53.

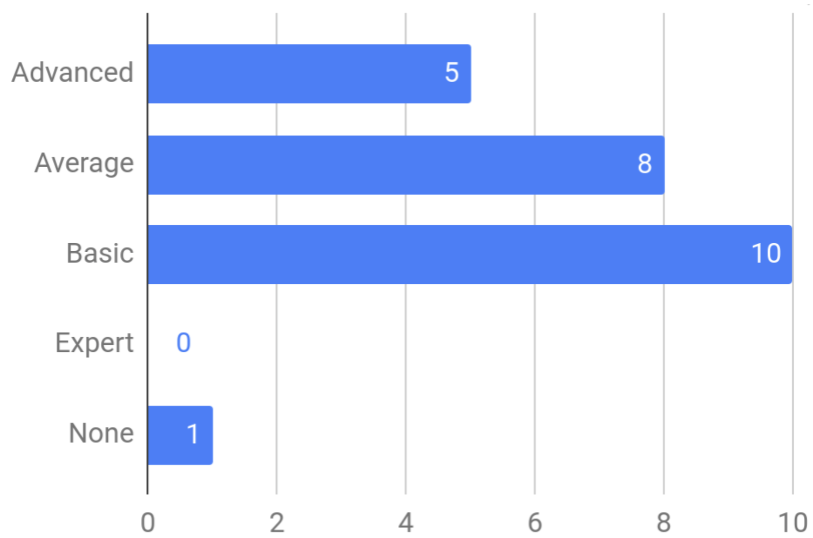

Figure 53. Knowledge of statistics

The questionnaire asked the participants about their experience with fantasy sports. Two thirds of participants (16/24) reported to have never played fantasy sports and 3 played rarely. None reported playing all the time or frequently and only 21% (5/24) of participants played occasionally. Among the 8 who reported some fantasy sports experience, 3 participants preferred fantasy football, cricket was the sport of choice for 2 of them, and 1 played fantasy soccer (Figure 54.) The remaining 2 did not report a preferred sport.

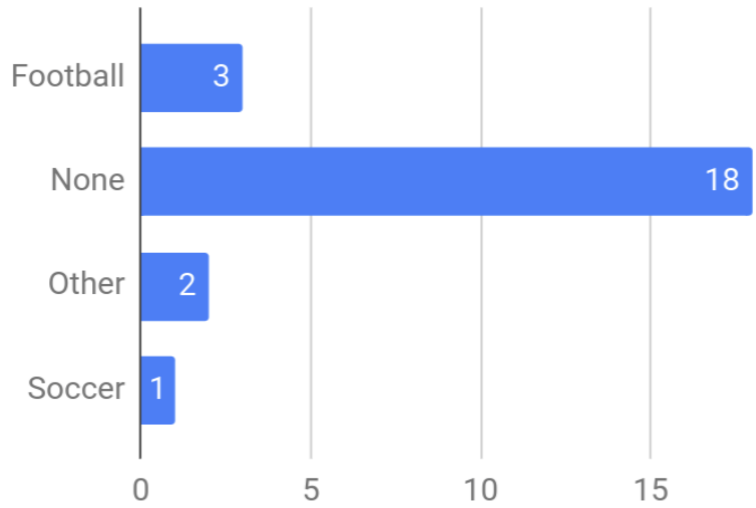

Figure 54. Sports preferences of participants who played fantasy sports

Participants also reported on how often they play online games other than fantasy sports (Figure 55). Only one participant responded to have never played games and no one said they played all the time. 5 answered that they play frequently while another 5 said they play rarely. Most respondents (13/24) reported playing games occasionally.

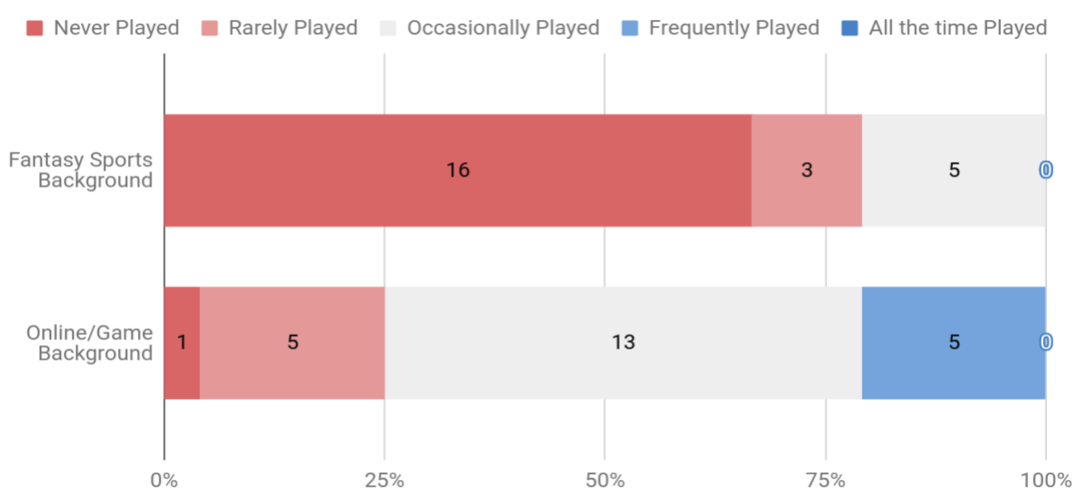

Figure 55. Game play frequency

In summary, the participants were mainly graduate students with a background in engineering or science with at least a basic knowledge of statistics. Although very few played fantasy sports, many were at least occasional game players. The following table (Table 3) maps every participant to their background to better understand the context of their comments. Each participant name is encoded using the convention {system version code} – {index}. For example, NT-A-1 is the first participant that used version 3 (No template – only automation) of the ACW.

Table 3. Participants and their background

Participant	**Background**	**Knowledge of Statistics**	**Fantasy Sports Background**	**Game Background**
T-A-1	Engineering	Average	Never Played	Frequently Played
T-A-2	Humanities	Basic	Occasionally Played	Occasionally Played
T-A-3	Business	Basic	Rarely Played	Occasionally Played
T-A-4	Engineering	Average	Rarely Played	Occasionally Played
T-A-5	Engineering	Advanced	Occasionally Played	Occasionally Played
T-A-6	Engineering	Advanced	Never Played	Occasionally Played
NT-A-1	Architecture Design	Basic	Occasionally Played	Frequently Played
NT-A-2	Other (Health Science)	Advanced	Never Played	Occasionally Played

Table 3. Continued.

Participant	Background	Knowledge of Statistics	Fantasy Sports Background	Game Background
NT-A-3	Physical Sciences	Basic	Never Played	Never Played
NT-A-4	Other (IT)	Average	Occasionally Played	Occasionally Played
NT-A-5	Social Sciences	Advanced	Never Played	Rarely Played
NT-A-6	Engineering	Average	Never Played	Frequently Played
T-NA-1	Engineering	Basic	Never Played	Frequently Played
T-NA-2	Physical Sciences	Advanced	Never Played	Rarely Played
T-NA-3	Engineering	Basic	Never Played	Rarely Played
T-NA-4	Social Sciences	Basic	Rarely Played	Occasionally Played
T-NA-5	Engineering	Average	Never Played	Rarely Played
T-NA-6	Engineering	Basic	Occasionally Played	Occasionally Played
NT-NA-1	Engineering	Average	Never Played	Frequently Played
NT-NA-2	Engineering	Average	Never Played	Occasionally Played
NT-NA-3	Engineering	Basic	Never Played	Occasionally Played
NT-NA-4	Business	Average	Never Played	Occasionally Played

Table 3. Continued.

Participant	Background	Knowledge of Statistics	Fantasy Sports Background	Game Background
NT-NA-5	Social Sciences	Basic	Never Played	Occasionally Played
NT-NA-6	Humanities	None	Never Played	Rarely Played

11. ACW USE: POST-QUESTIONNAIRE AND TASK DURATION RESULTS

11.1. Post-questionnaire

The first task performed by participants was the authoring of a prediction game activity within the context of the weather data collected for Fantasy Climate using the ACW. After the participants completed the task, they answered another questionnaire (Appendix I) collecting data about the participants' authoring experience and examining whether the template component and the automated components have an effect on this experiences. The questionnaire included 20 5-point Likert-scale questions. The first 10 questions were to measure the usability of the ACW using the standard system usability scale (SUS) [51,60]. They were followed by 2 questions about the template, 4 questions on the explanation system, 2 questions on creating the prediction schedule, and another 2 questions on creating the selection sets. We found no difference in the responses between participants of the four versions of the system. Thus, in the following paragraphs, we only report on overall results.

The SUS score of the ACW was 64.38 which is below the average: 68 [51,52,60]. SUS scores are like grades (for a test or a quiz). As a consequence the ACW's usability grade is a D [51,52,55].

The template component was in two versions of the ACW hence available to 12 participants. 10 of them agreed that the template was helpful in their task and the other 2 were neutral. To the question whether the template was unnecessary to author the prediction activity, only two participants were neutral. 5 at least disagreed and the other 5 at least agreed. Figure 56 illustrates.

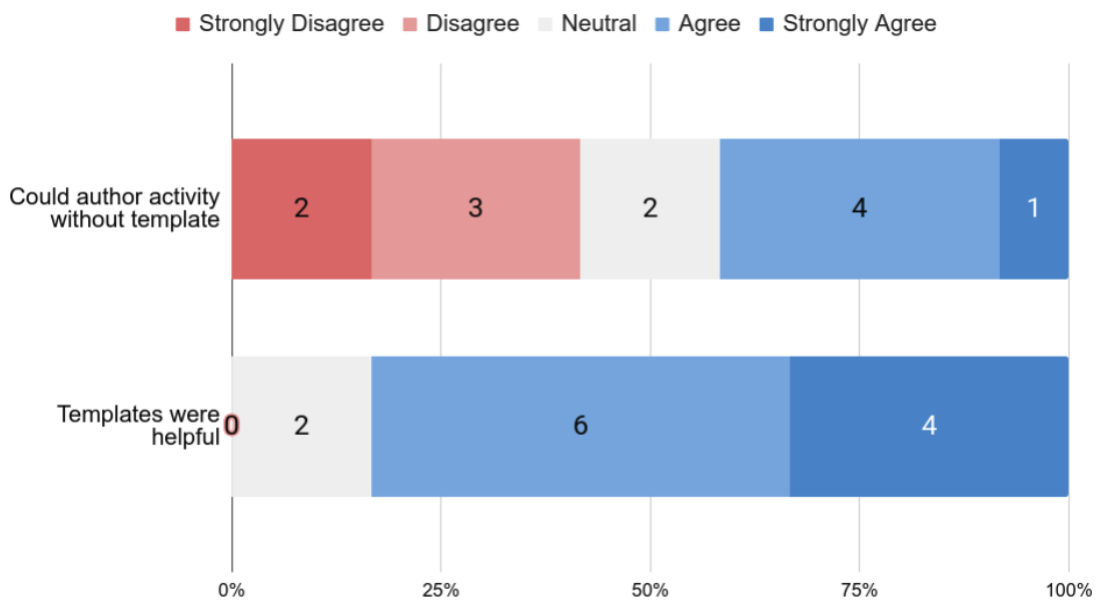

Figure 56. Sentiments on the template component of the ACW

The explanation system was constant in all four versions of the ACW. Figure 57 shows that almost all the participants (22/24) agreed that the help signifiers (the clickable question mark icons on every interface of the ACW) were easily visible. Also, about 2/3 (17/24) agreed that the explanation were easy to read and understand. On the question of whether the automatic popup of help at the beginning of every step was annoying, there was a split in the responses with 9 participants agreeing and 10 disagreeing. 5 were neutral. And, only one participant agreed that the explanations were not helpful during the authoring task.

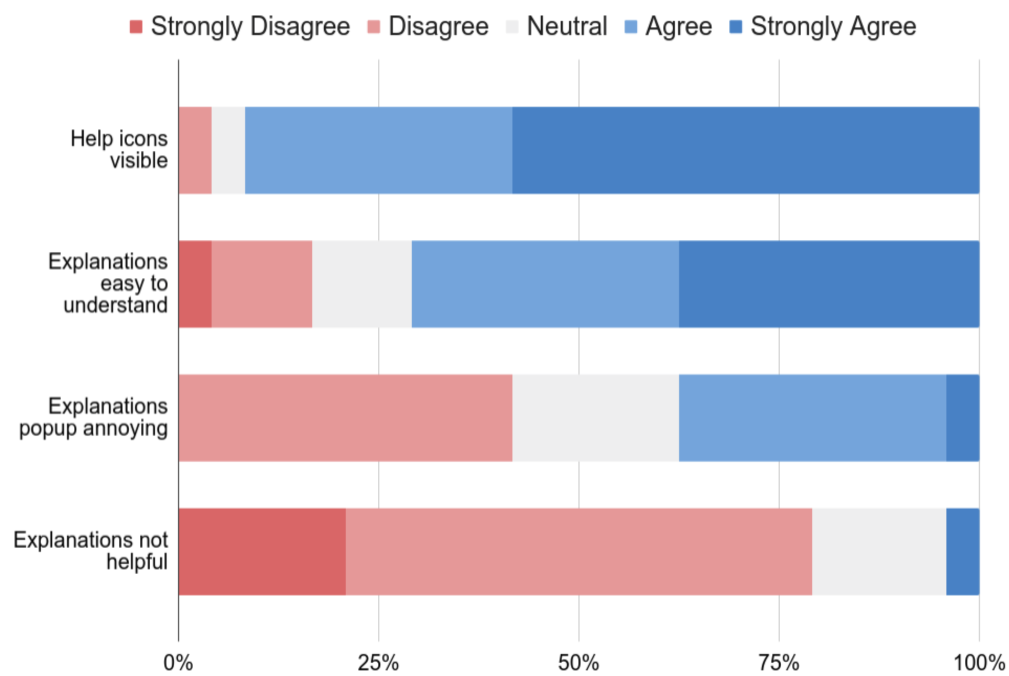

Figure 57. Sentiments on the help system of the ACW

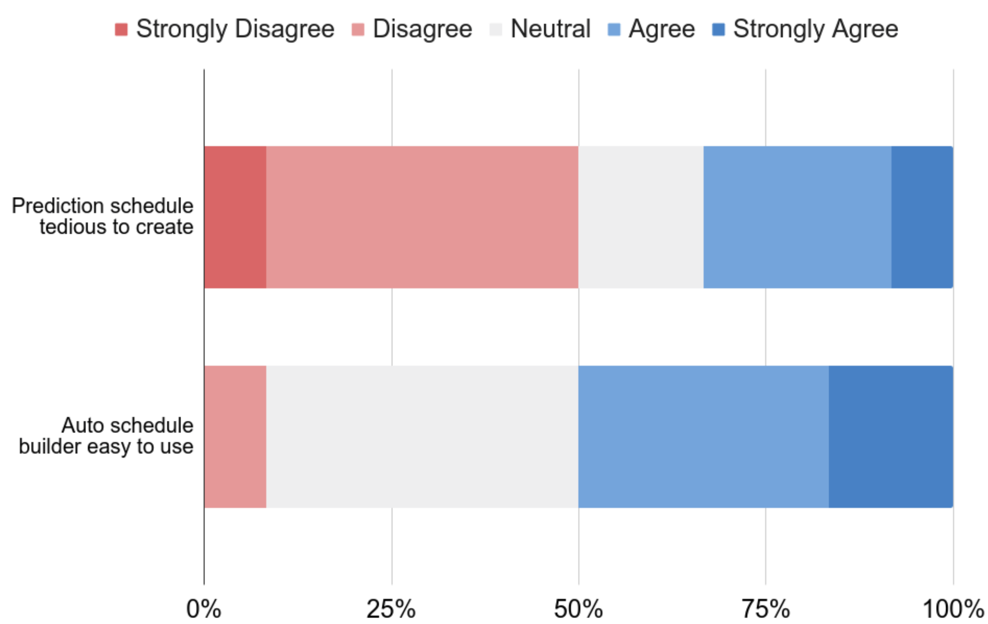

Figure 58. Sentiments on creating the prediction schedule and using the automatic schedule builder

When asked whether creating the prediction schedule was tedious, half of the participants (12/24) disagreed, a third agreed (8/24) and, 5 were neutral (Figure 58.) The automatic schedule builder was only available in two version of the ACW hence to 12 participants. Only 1 of these 12 participants disagreed that the automatic schedule builder was easy to use. Among the rest (11/12), 5 were neutral and 6 agreed.

Regarding the creation of selection sets for the prediction activity, 14 out of 24 participants agreed that it was easy, 9 were neutral, and 1 participant disagreed (Figure 59.) Like with the automatic schedule builder, only 12 participants had access to the selection sets generator. 75% (9/12) disagreed the selection sets generator was difficult to use and 1 agreed. The rest were neutral.

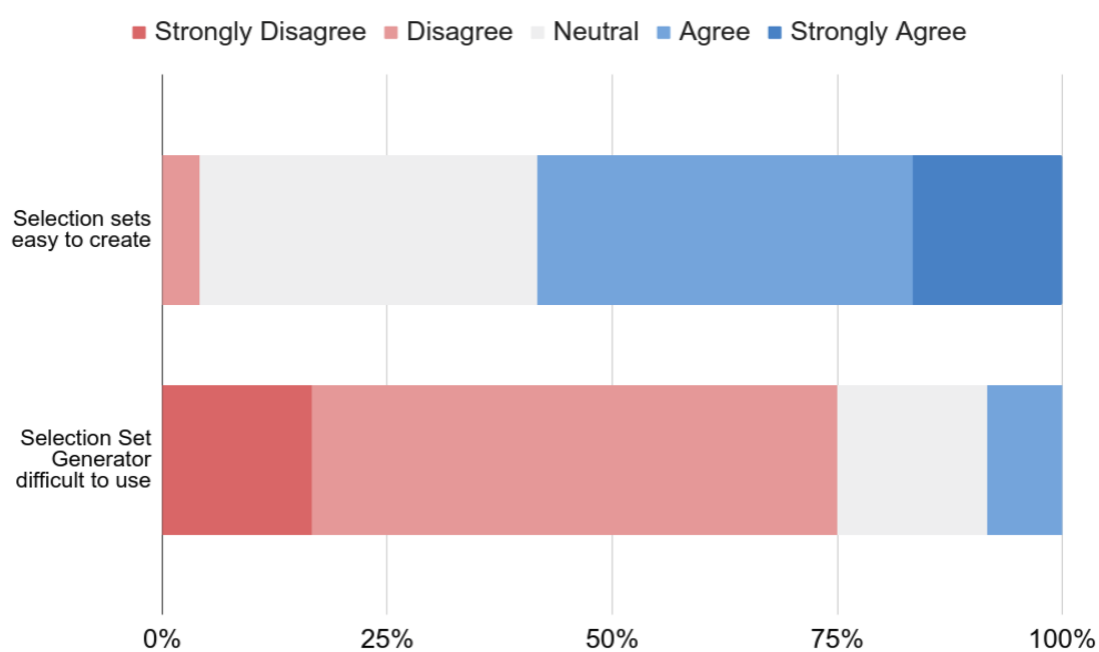

Figure 59. Sentiments on creating the selection sets and using the selection sets

The above responses give a sense that many participants found the prediction activity authoring to be complex but manageable. The templates and explanations were predominantly viewed as valuable while the automatic generation of schedules and selection sets was viewed as beneficial by about half of the participants with versions of the ACW providing these modes of support. There might be many reasons why the SUS score was poor. SUS is “not intended to diagnose usability problems” and its “results won’t shed much light on why users are responding the way they are” [51]. What participants found difficult is not exposed by these results. In addition, “difficult tasks lower SUS scores by 8% on average” [52]. The authoring of a prediction game, something previously only done via programming, is most likely such a difficult task. Where participants spent their time in the authoring process may help identify where participants were most challenged. The next section explores this aspect of the data. The rest of the results from the questionnaire may be better understood based on the interviews conducted after participants completed the authoring task. Later sections describes the results from those interviews.

11.2. Task Duration

Using the screen capture videos of the participants’ interaction with the ACW during the authoring task, we extracted the overall duration of the authoring process and the duration for every step in the ACW. The average duration of the authoring task was 1736.37 seconds (stdev = 708.74) or just under 29 minutes. Figure 60 shows the average duration (in seconds) for each step of the ACW. The longest steps were Step 2: making the prediction schedule, Step 3: creating the selection sets, and Step 4: creating the scoring formulas.

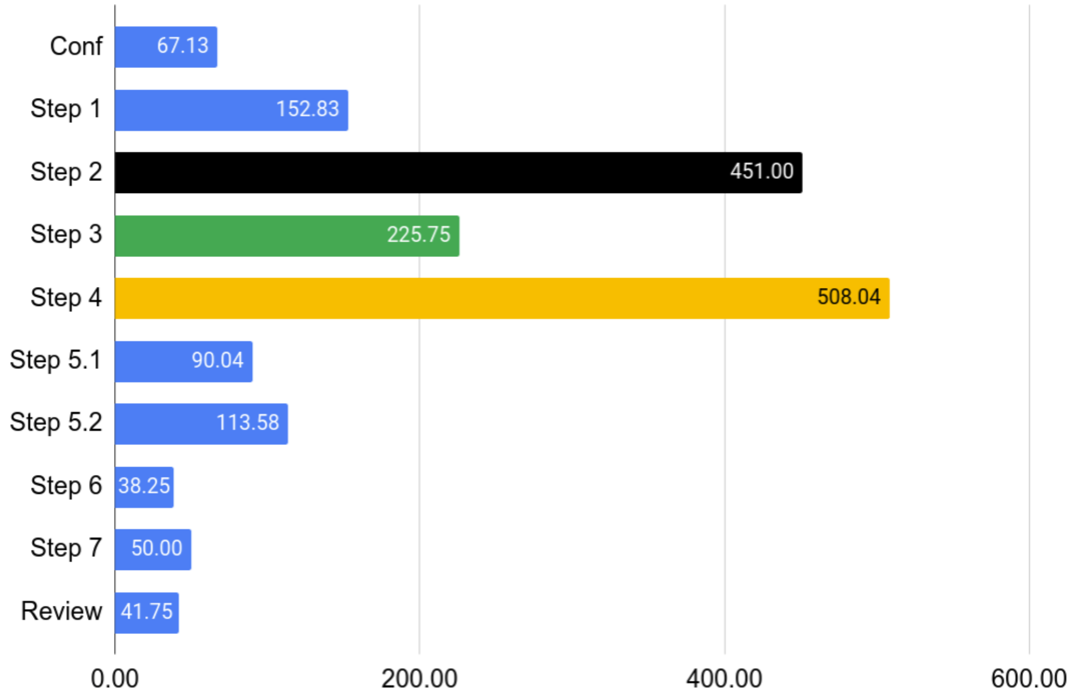

Figure 60. Average time spent (in seconds) for every step of the ACW

Not all participants with access to the template or automation capabilities made use of them. The following comparisons examine the effects of use of these features on the authoring process, so only include those participants that actually used a capability. Later we discuss the rationale of participants' use or lack thereof.

Using a template to author a prediction activity reduced the time of the authoring process. As shown in Figure 61, the participants who used the template spent less time than the ones who did not use the template. Similarly, automating repetitive aspects of the authoring process should also help with the time required for those steps. Figure 61 shows the average time to complete steps 2 and 3 for those who used automation versus those who did not.

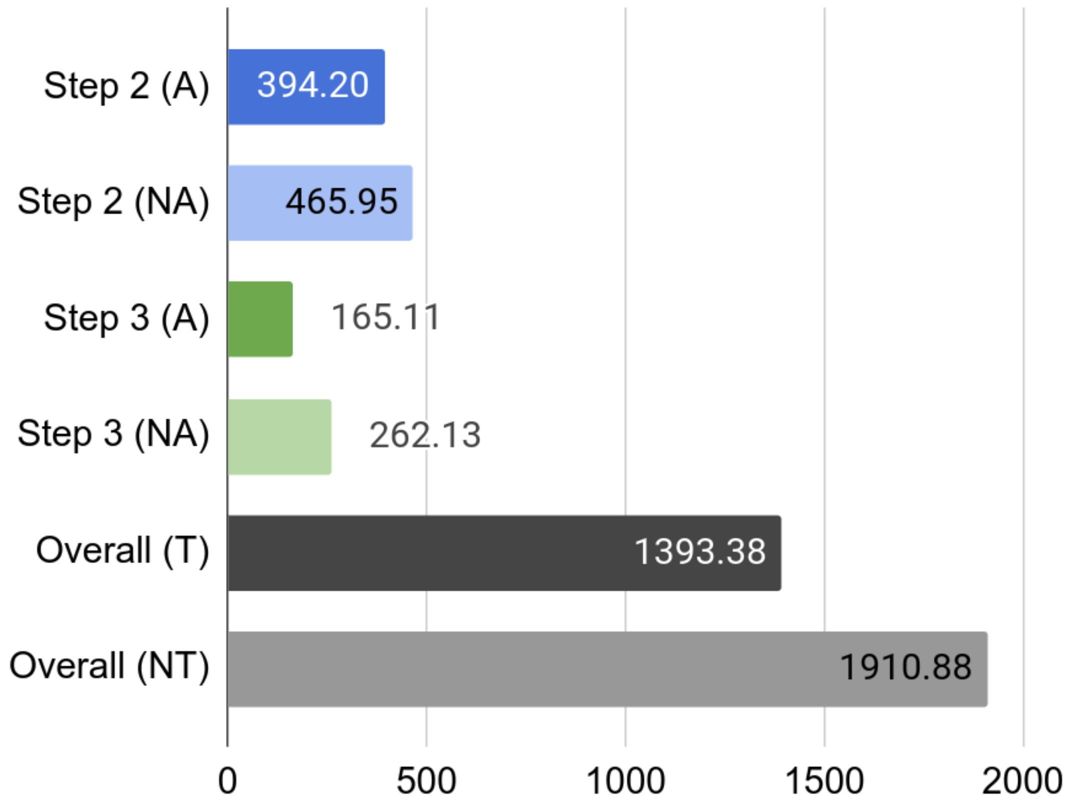

Figure 61. Step 2 and Step 3 average duration (in seconds) with automation (A) and without automation (NA) and overall average duration with template (T) and without template (NT)

12. ACW USE: INTERVIEW RESULTS

Following the Likert-scale survey was a short semi-structured interview (Appendix J) aiming to understand perceptions of the authoring task and experiences with the ACW. The participants' responses were transcribed from the audio recording. Quotes from participants include square brackets to provide more context for the participant's answer. As previously described, the convention {version code} – {index} is used to identify participants. For example, T-NA-3 is participant 3 that used the ACM with the template but without automation.

12.1. Results

12.1.1. Overall Experience

The interviews began by asking the participants about their overall experience using the ACW to author a prediction activity for Fantasy Climate. Among the 21 respondents, 71% (15/21) described their experience as positive using words such as 'easy' (11/17), 'good', or 'cool': T-NA-2: "*I think it was pretty good. Very positive...it was easy to navigate through. Instructions were clear,*" NT-A-1: "*It was really easy to use....I felt like it was like basically holding your hand through the whole thing*" and, NT-A-6: "*It was pretty cool. Pretty cool.*" 2 participants (out of 21) expressed a neutral view on their overall experience because the domain (climate/weather) was uninteresting to them: T-A-1: "*[Researcher: the domain weather is not interesting to you?] Yes!*" and, T-NA-4: "*I could also see it being pretty cool if I was creating it for something a little different or something that I have more interest in personally*." Another 2 (out of 21) explained that, without the proper context (i.e. Fantasy Climate demonstration done at the beginning of the experiment) or a template activity as a starter for customization, using the ACW would have been difficult. Four participants reported a negative experience authoring the prediction activity. Two of them viewed the task as tedious: T-A-1: "*…this interface it seems like to submit a CV. It's [like] applying for an interview. It seems like working, not [a] game.*" The other two thought the task was complicated: NT-NA-5: "*It was complicated. Especially like it being the first time. [this is] something I would not wanna do on a daily basis.*"

The following table (Table 4) identifies the participants' overall sentiments (P=positive, U=uninteresting domain, C=conditionally difficult, N=negative) with the version of the software used and to their use of the template (TPL), the automatic scheduler builder (ASB), and the selection sets generator (SSG). For instance, T-A-1 had access to the template, ASB, and SSG but used only the SSG. This participant mentioned the domain was uninteresting, aspects of authoring were conditionally difficult, and had an overall negative sentiment about the process. Gray cells indicate the participant did not have the option to use a tool due to their assigned software configuration.

Table 4. Participants' overall sentiments mapped to their use of the template, the automatic schedule builder, or the selection set generator

Participant	P	U	C	N	TPL	ASB	SSG
T-A-1		X	X	X			X
T-A-2	X						X
T-A-3	X				X		
T-A-4			X		X	X	

Table 4. Continued.

Participant	P	U	C	N	TPL	ASB	SSG
T-A-5	X				X		X
T-A-6	X				X	X	X
NT-A-1	X					X	X
NT-A-2	X					X	X
NT-A-3	X						X
NT-A-4	X						
NT-A-5						X	X
NT-A-6	X						X
T-NA-1	X						
T-NA-2	X				X		
T-NA-3					X		
T-NA-4		X					
T-NA-5					X		
T-NA-6				X	X		
NT-NA-1	X						
NT-NA-2				X			
NT-NA-3	X						
NT-NA-4	X						
NT-NA-5				X			
NT-NA-6	X						

12.1.2. Required Background

Was a background in games or fantasy sports required to be able to use the ACW? Only 3 out of 23 participants agreed that a background in games was required. One of them explained that the reason may be because gamers are better computer users: NT-NA-2: "*...so for a gaming background, I guess they know different techniques, they're aware of different techniques.*" However, the overwhelming majority (87%, 20/23) said no gaming background was necessary: NT-A-5: "*Not at all. I mean, I do not play video games or anything of that sort, but it was pretty [self-explanatory],*" T-NA-2: "*No, I don't think so. It did not really feel like a game*" and, NT-NA-3: "*you or someone without a background can do it after learning.*"

Regarding a fantasy sports background, only half (12) of the participants answered since the other half did not have any experience playing fantasy sports. For most respondents (11/12), no background in fantasy sports was required to be able to use the ACW: T-NA-1: "*I don't think creating this game was very related to actually playing it. So the experience was too far from that.*" In these, two (2/11) explained that although experience with fantasy sports was unnecessary, having one would help: NT-NA-3: "*No, I don't think you need that background. But like if you have that background, it will make it easier.*" Another participant (1/11) elaborated that, instead, a background and interest in the domain (climate/weather) was necessary to use the ACW: T-NA-4: "*the actual climate aspect...is what makes it a little challenging in terms of...those determinations...Most people that play fantasy sports have...some love of the game. So you would have to find somebody who is very...interested in weather patterns...*" Only one respondent (1/12) affirmed that using the ACW required a background in fantasy sports, especially for the formula step.

12.1.3. Help and Explanations

Explanations provided detailed information (context, rationales, and examples) on every aspect of the ACW user interface including the effects on the prediction activity in Fantasy Climate. At the beginning of every step during the authoring process of a prediction activity, dialogs containing explanations would automatically pop up to guide in the customization (Figure 62). 18 participants elaborated on their experience with the automatic explanation popups. 67% (12/18) found the popups to be bothersome, with 7 of them expressing that they preferred to see or try the interface first: T-A-1: "*it will show up...explaining first. But...I don't know what is the interface...maybe I can just...see the interface first...*" and, T-NA-3: "*I don't want to start something with the pop up...Because first I want to figure out what it is on the screen.*" The other 5 qualified their experience using words like 'distracting', 'difficult', or 'annoying': T-A-5: "*The only the problem I faced was the popups, they were annoying*" and, T-A-2: "*I find [it] a bit...distracting...when you select something you have to exit...I was thinking if I exit I will be able to destroy it. I thought it a bit difficult.*" In the latter quote, the participant was referring to the red 'Exit' button (Figure 62) on the popup dialog. Although pressing that button closed the explanation dialog, s/he interpreted it as exiting the ACW and losing all prior customizations.

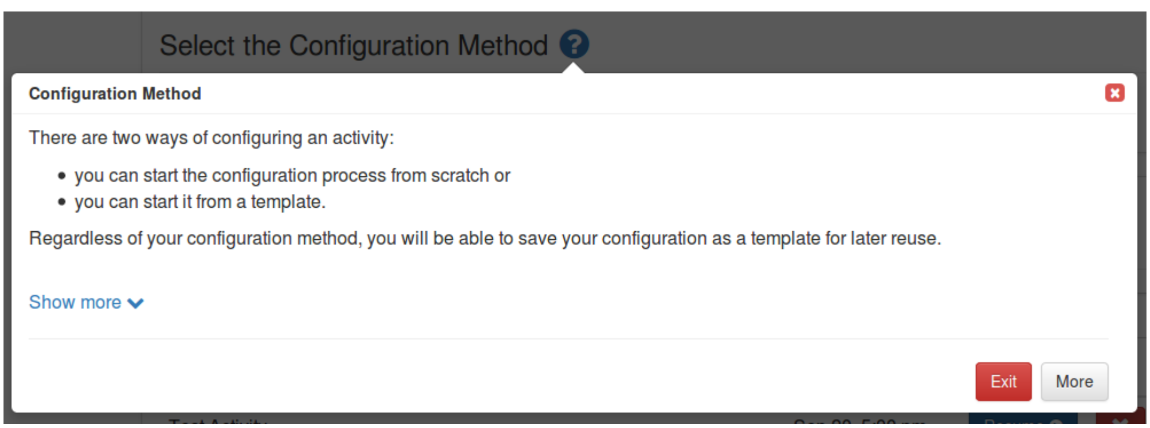

Figure 62. Explanation dialog automatically popping at the beginning of the step

According to 4 out of 18 participants, the automatic explanation popups were either useful (3/4) or unnecessary (1/4) in certain circumstances: T-NA-3: “*it might be friendly for the user who is doing it for the first time...I would not want to see them every time I customize...*” or, T-NA-1: “*I thought in certain places it wasn't [required]...it's very obvious to [click on it and understand it...[it was required] just for the formula section because it wasn't easy to...figuring out [by ourselves]*.” But 5 out of 18 participants thought that explanations dialogs automatically popping up at the beginning of every step in the authoring process was helpful: NT-NA-1: “*...It was very helpful.. I think it [would] have been very difficult without...the dialog popping up, showing you what to do and how to do it*” and, NT-NA-4: “*Those are really helpful for me...I liked the fact that it would pop up cause it would just...go ahead and...provide the instructions of exactly what I'm supposed to do.*”

Besides the automatic popups, participants could also request explanations if needed by clicking one of the help signifiers (question mark icons) on the interface (see any figure in chapter 9.) But did these explanations helped the participants accomplish their task of authoring the prediction activity? Among 20 responses, 70% (14/20) confirmed that they did: T-A-6: “*Yeah...I was able to see those examples...it was...far quicker to do it rather than like trying to navigate on my own without any example*”, and T-NA-1: “*Yeah. Otherwise you'd have to ask somebody [...what] this function does, I think that helps.*” For 2 participants out of 20, explanations were only useful in the formula creation step of the ACW. Another 2 thought they were helpful except in the selection sets step. The explanations were not helpful to the last 2 participants: T-A-5: "*...not reading at all because...I was understanding what was going on.*"

Half of the participants (12/24) further elaborated on the content of the explanations. Two of these expanded on the utility of explanations: NT-A-4: "*...those were quite helpful. I mean...the explanations were short...it was helpful in simple language and more importantly, there were examples for each explanation...which is important because...it's quite detailed and...quite a process. So examples [did help] so yeah...that was [a] very helpful feature.*" One participant highlighted unfamiliar or unclear language in the explanation: T-NA-4: "*It was just some of the language that it uses just takes a second to...read through...[I] had to read it...four or five times and be like okay. I think I get what this is asking you to do...I think one of them that probably...bugged me [and], kind of got to me was the 'persist the result.' [It] is [the] example where it said, and if you click this then it can 'persist'. I was just kinda like okay...To me, persistence is struggling or persevering through something.*" However, most comments (11/12) were about how the explanations were overly detailed and very wordy: NT-NA-2: "*I think that the explanations are so detailed that I am having trouble to even follow...after a certain point of time, I am not feeling like reading so much*", and NT-NA-5: "*I feel like they're very wordy. Yeah. Like information overload, you know?...a lot of...unnecessary info cause I would read it and be like, okay, I dunno. It's just complicated.*"

A third of the respondents (8/24) provided feedback and suggestions for the explanations. One of them advocated for disabling the automatic popups because explanations accessible on demand by clicking the help icons on the interface was sufficient: T-A-5: "*disable the popup. If people wanted more information...they could just go on and...get more information...clicking on the question mark.*" Four recommend that all explanation content should be in a constant panel adjacent to the ACW main interface as Figure 63 illustrates. Two preferred the explanations to have interactive or media rich content: T-A-2: "*I would rather prefer having a video.*" And one participant suggested to make the explanations content searchable through a user interface component (e.g. search bar) in the ACW.

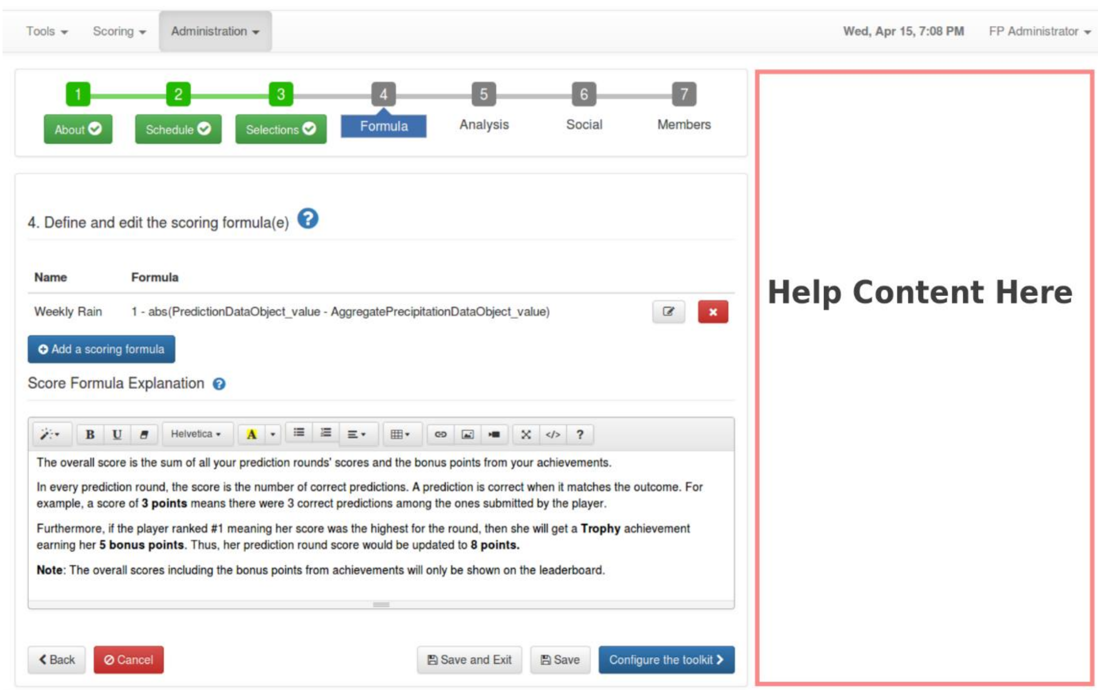

Figure 63. Explanation panel on every ACW interface

12.1.4. Template Use

The template component was available to the 12 participants in versions T-A and T-NA. 66% (8/12) of them made use of the template. Among these, all except one updated the presets of the template. For example, 6 updated the preset scoring formula whereas 4 updated the preset selection sets. Three of the four participants with access to the template that did not use it were asked why? One stated they did not see the button to the start the ACW from a template. The other two interpreted the study task's instructions as configuring the activity from scratch: T-NA-4: "*[I] was thinking of it in terms of I had to create my own configuration. So going forward from that way*."

12.1.5. Step Order in Authoring Task

The ACW enabled the authoring of a prediction activity through a sequence of steps. We asked participants whether the current order of ACW steps made sense or whether they would have preferred a different order. 21 answered. All except 2 (90%) affirmed that the authoring process sequence of step made sense: T-NA-6: "*it's the perfect sequence,*" NT-A-2: "*it's easy to follow. I think it's logical,*" and T-A-6: "*I felt it was orderly...this was perfectly perfect.*" Of the two participants that preferred a different order, one commented that s/he would had favored working on the formula step and the analysis step (choosing what analysis tools should be available to the players during the prediction – Figure 41) last because they were harder than other steps. The other participant elaborated further on his preference for configuring the selection sets (ACW Step 3) before creating the prediction schedule (ACW Step 2): T-NA-4: "*I could see doing the selections before scheduling and then having the location [sets] built into the scheduling aspect. So then saying like on this week you're going to go with location set one, on this [other] week you gonna go with location set two, and that way you can* *finagle it because on this level you kind of don't have any control of what location set goes with what schedule.*"

12.1.6. Prediction Schedule Configuration

In step 2 of the ACW, participants made a prediction schedule for their activity by setting the activity start date, the activity end date, the prediction due dates and the prediction end dates. For one third of the participants (8/24), this step of the activity authoring process was the most or one of the most difficult. For 5 of these 8 participants, making the prediction schedule was the most or one of the most difficult steps because it was hard or tedious: NT-NA-4: "*I had to go through and find...the exact date for...every single one and be like....this is two weeks so far...I think...this sounds...so tedious ...clicking...through...the calendar...over and over again...I kind of would get lost a couple times.*" Among these 5 participants, only 1 used the automatic schedule builder. The rest did not because they were in versions of the ACW with no automation (T-NA and NT-NA).

Another reason (2/8 responses) was due to unfamiliar or unclear terms on the user interface of the automatic schedule builder (ASB): NT-A-1: "*It just wasn't super clear to me. Like, I don't know what [it] meant by 'last round offset.' Like 'prediction round' and 'prediction round interval' I got but then this one [last round offset], I was like, Hmm. I'm not quite sure....It's just the wording of it.*" The final reason (1/8 response) was considering the game in a peer-use setting, where creating the prediction schedule required the consent of others that would partake in the future activity: T-A-1: "*I should also choose a person I want to compete with and if I chose a date by myself,..I* *need to communicate with that person. The person I choose maybe, I don't know. So I don't want to talk...with the person...about the exact dates: the due dates and the score dates.*"

16 participants for whom ACW Step 2 was not the one of the most difficult elaborated on their experience. About half of them (7/16) corroborated that making the prediction schedule was tedious: T-A-5: "*The schedule was very cumbersome to make*" or, NT-A-5: "*it was actually tiring, difficult because [of] keeping track of all the dates.*" Nearly all of these participants (6/7) did not use the automatic schedule builder because some (3) were in versions of the ACW with no automation and the others (3) who had access to automation chose not to use it for reasons elaborated later.

Four participants brought up a couple of issues with the user interface. One was paging: NT-A-3: "*...after five rounds, [it] went to the second page. But I did not realize that there is a second page. Yeah. That's confusing...I [thought] everything got deleted somehow.*" On the prediction schedule making interface, only five prediction rounds (due dates and score dates) are displayed at one time. For bigger schedules, the system creates pages showing at most five prediction rounds. Then the user can navigate between pages to view or edit a prediction round. In Figure 64, the prediction schedule contains six prediction rounds. The top shows page 1 and the bottom shows page 2. The other issue regarded the user interface component that let the author input a date and time: T-NA-4: "*...for some reason if you hit delete (on the keyboard) it actually wipes the entire window...I kept getting confused [because] I used the delete button (on the keyboard) to delete...things...I think it got me like three times.*" This comment was referring to the fact that while editing a date/time, if one pressed the 'Delete' key to delete a single character, the date/time component would instead clear the whole input (Figure 65.)

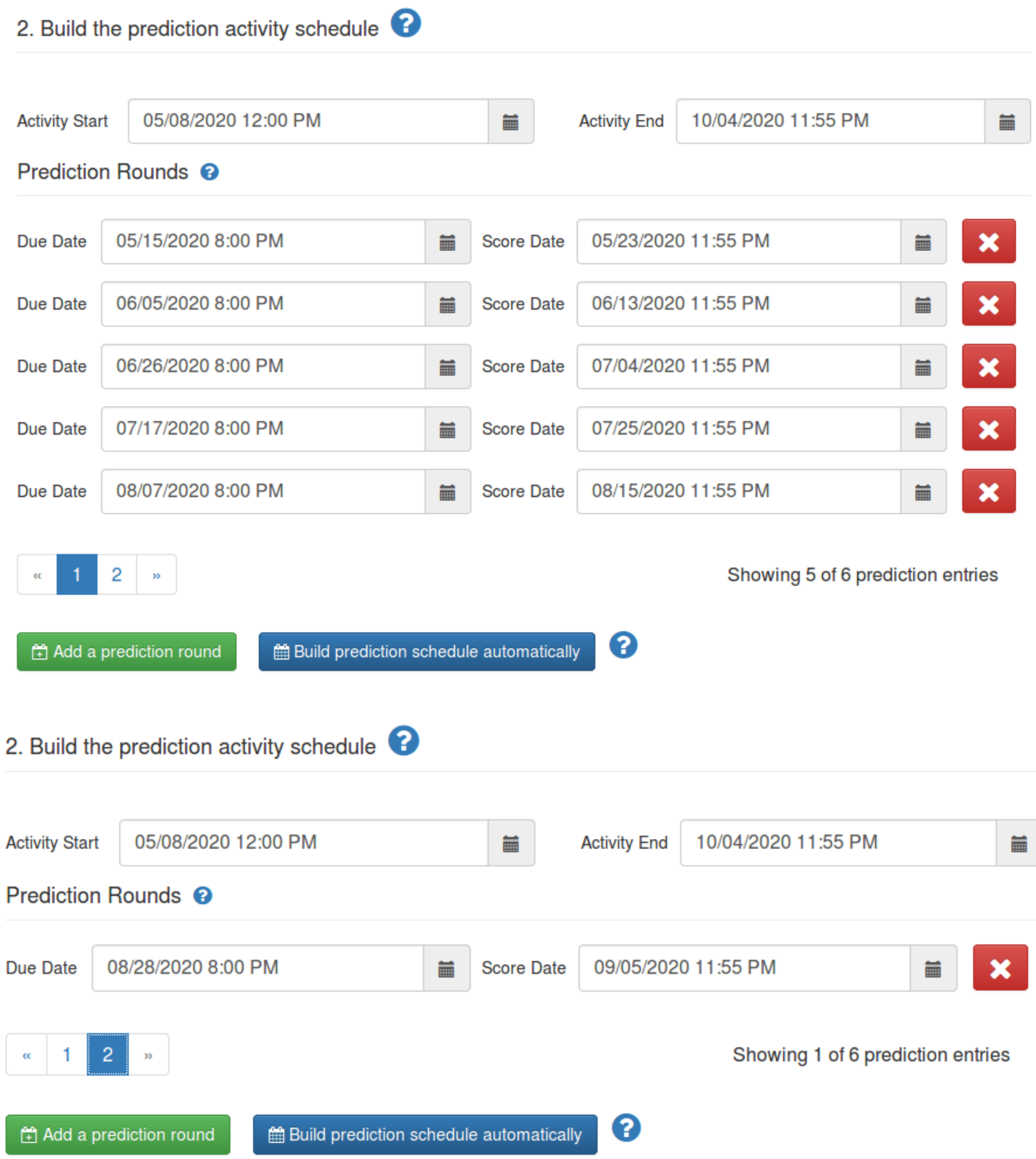

Figure 64. A prediction schedule containing six prediction round. The top image shows page 1 of the rounds and the bottom shows page 2

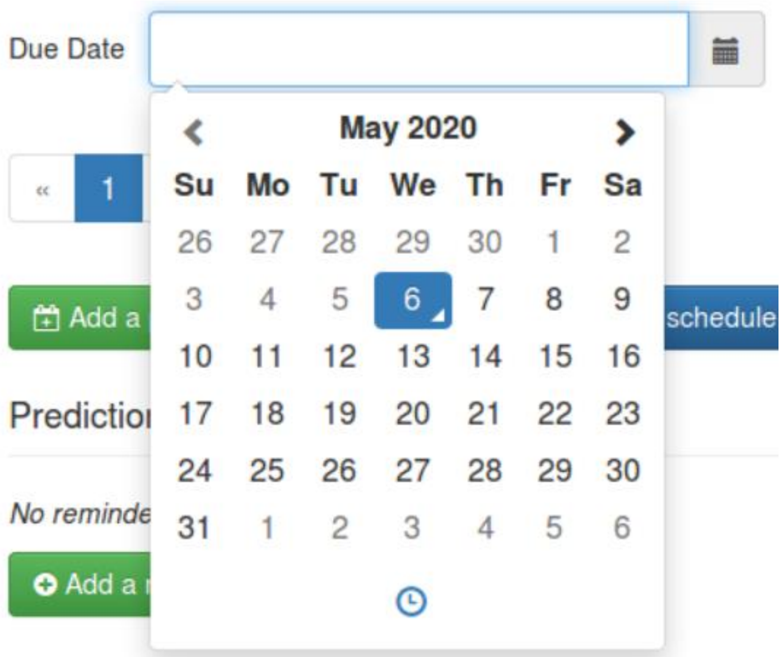

Figure 65. Date/time component clears user input when the 'Del' button is pressed on the keyboard

Nevertheless, making the prediction schedule was a positive experience for several participants (9/16). 3 of them qualified it as good: T-A-3: "*so it was good...it's just simple enough.*" For one participant, the automatic schedule builder made it easy: NT-A-2: "*I [needed] to go in there and do it one by one, but the automatic scheduling is really helpful. So easy...it actually corrected my end activity date...I didn't have to spend the time counting the weeks and trying to calculate et cetera. It [was] quicker and easier.*" And 5 reported that making the prediction schedule was straightforward: NT-A-4: "*that was quite straightforward given the problem statement of being 14 weeks and then two weeks and then also clarify about the Friday and the Sunday. So yeah, that was very easy to use.*" Only 1 of these 9 participants used the ASB even though 4 of them were in versions of the ACW where it was available (T-A, NT-A). The others (5) had no access to automation in their versions of the ACW.

Furthermore, half of the participants (12/24) suggested improvements for the prediction schedule making step of the ACW. One was that in an informal setting when one is authoring an activity to be played with strangers, the system should automatically generate the prediction schedule. Another suggestion, from participants (7/12) who did not use any automated tool, was for a tool very similar to the automatic schedule builder: T-A-5: "*If you can make an automatic wizard like you made for do this, enter the two dates and they will generate it automatically two weeks or three weeks or whatever you want. You [can] add more conditions to it. If you want weekly then it could generate weekly and every day...make it automatic.*" The last suggestions were about user interface improvements such as making the button that launches the Automatic Schedule Builder more visible (1/12), reducing the cognitive effort to track dates (1/12), and improving the date/time inputs (2). For the last user interface improvement, one participant specified that prediction due time and prediction score time be set only once since they will be constant during the activity: T-NA-4: "*I could also see the advantage of taking the time out of the prediction...I can see it saying like...scores and due are at five and then putting like due date scoring date and not having to put the time, cause I feel like the time is almost always going to match.*" Another one proposed to split the date/time component into two components: one for the date and another for the time. Figure 66 illustrates.

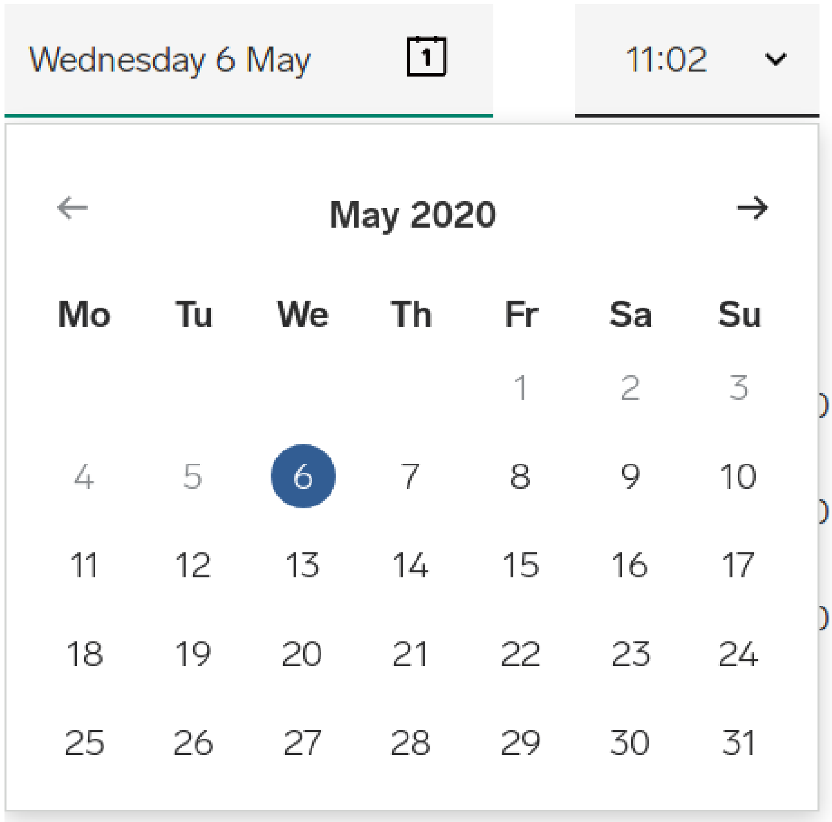

Figure 66. A date/time component where the date input is separated from the time input

Even though the Automatic Schedule Builder (ASB) was available to 12 participants in versions T-A and NT-A of the ACW, only 5 actually availed themselves of it. The others (6) explained why they did not use it. One respondent explained that s/he thought the instructions required that the prediction schedule be built manually. Another one reported that s/he did not see the button that launched the automatic schedule builder. 2 respondents opined that automated tools were generally restrictive: NT-A-4: "*My experience with...the auto generators is that...there's a limit to customization... You may or may not get [what you want], I could have of course opened* *[the ASB] and checked it and it might have...been sufficient. But...quite often...it's not enough.*" And the workings of the ASB were unclear to half of them (3/6): NT-A-3: "*when I went here [the ASB interface], I thought this build something wrongly...I did not...pay attention to this [prediction round interval and last round offset components of the user interface]...I would have gone for it if I saw it more carefully and understood how it worked, but I didn't pay attention and just went back...This [ASB] is easier.*"

12.1.7. Generating Selection Sets

The third step during the activity authoring process involved creating the selection sets: the sets against which the players would make predictions. In Fantasy Climate, these selection sets are composed of the locations of U.S. cities. This third step of the ACW was one of the most difficult for 5 participants out of 24. There were two reasons. The first one coming from 2 participants was because of unknown system requirements: NT-NA-2: "*I did not know that I have to create three sets of minimum five location in each.*" The problem was that the ACW required that any activity created for Fantasy Climate must have at least three selections to make the game less repetitive and avoid players making predictions against the same sets. The other reason coming from 3 participants was due to user interface issues. One of them said: T-NA-5: "*I didn't see a direct link between this [map visualizing the locations] and this [location sets list/table below the map.] That's the issue...So I didn't think that once you fill out this [set of locations by creating the list right adjacent to the map] right, you can add it to here [the table below the map]*." Figure 67 illustrates. The other two were confused that the option to create and configure was hidden inside a dropdown component (as shown in Figure 68) hence not easily visible by the user.

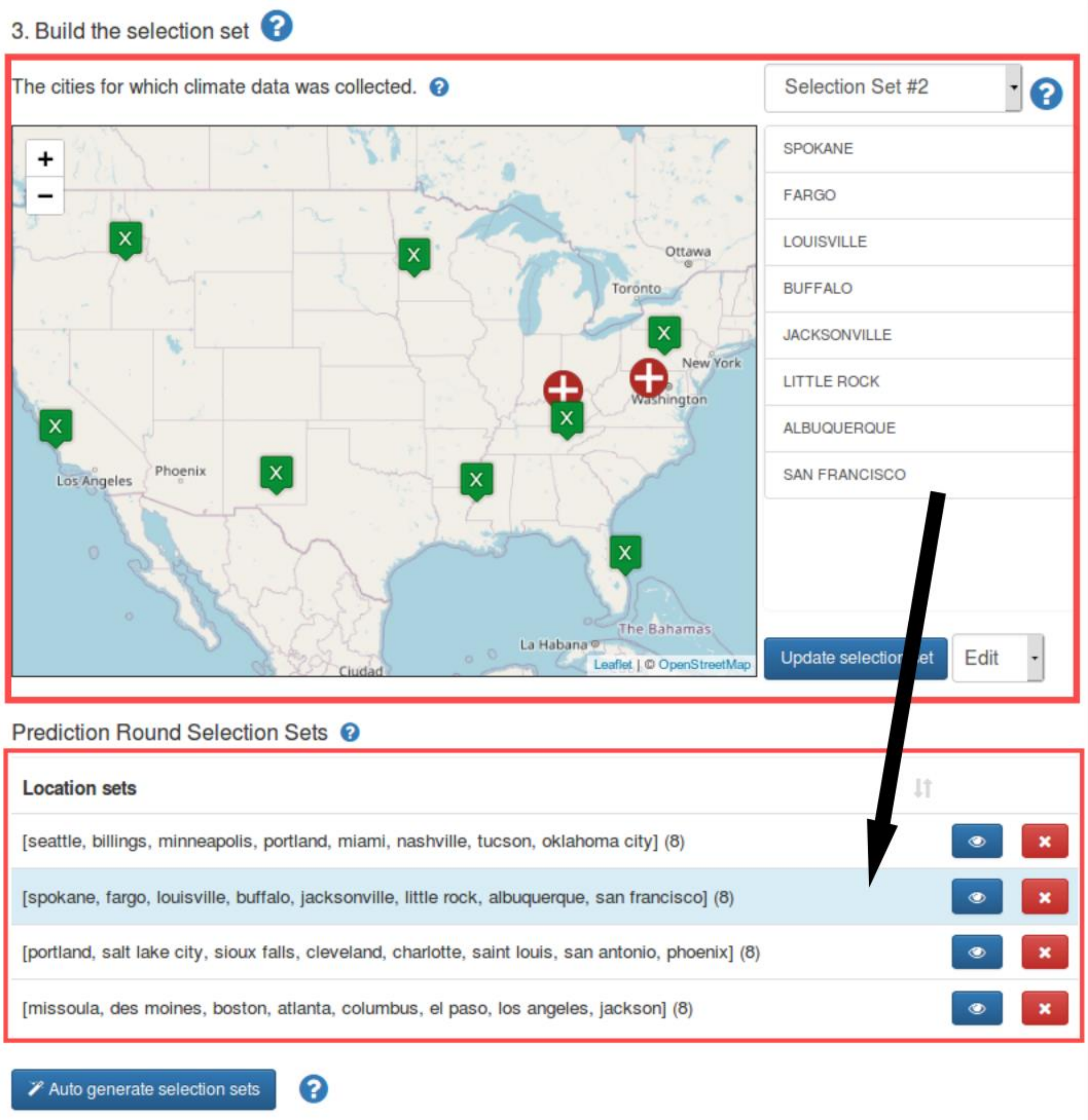

Figure 67. Selection sets interface – highlighting the locations set being created from the map and the list of already created location sets

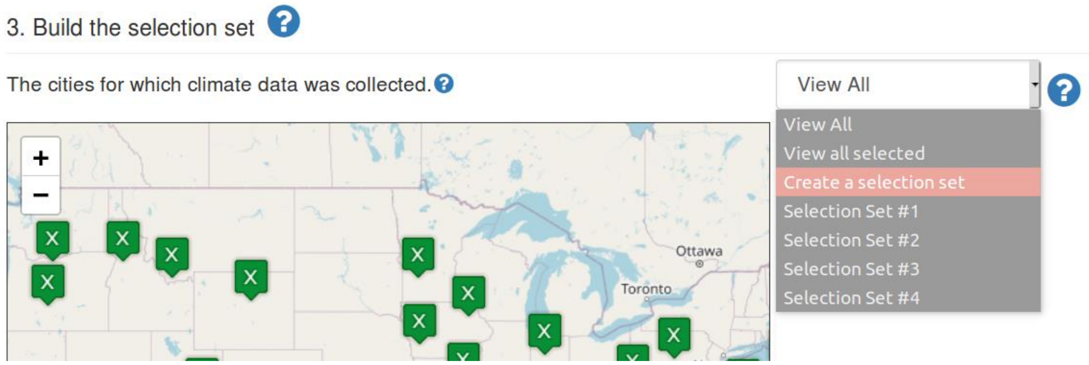

Figure 68. The option to create a selection set hidden under a dropdown

Additional insights about challenges encountered generating selection sets can be found in the comments of the 19 participants for whom this was not the most challenging step or even challenging for them. The ACW requirement and its rationale that a minimum of three selection sets be created was also obscure to 5 of them: T-NA-1: "*You have to create at least three sets. So I wasn't sure what that was for.*" A few (4/19) brought up more user interface issues. They especially highlighted that the signifier icons used on the map part of the interface were at times misleading: NT-A-2: "*I thought that the plus means I can click and add, but I can't, I guess you need to go into the edit mode, which is...not available directly here. And it's...hard to find...those little crosses [the 'x' on the map location marker.] I assume they're like remove and I try [to] remove again, I couldn't do anything so I was kind of lost. I can't add, I can't remove*". Some other participants (4/19) explained why they used presets. One used the preset from the template because: T-NA-3: "*I'm from India, I don't know which city in the US are the...coolest...if I had a fair idea of what cities can be cool...I would have [created] my own custom sets.*" Three participants used the presets created by the selection set generator because it was quick: T-A-6: "*I didn't use the manual one. Probably that would have been a little more time consuming, but with this [autogenerator, it] was quicker*", and because it was easy: NT-A-2: "*First of all, what am I supposed to do? Kind of scary to me. I just went for the auto generator. That's easy and helpful.*" Nevertheless, 42% (8/19) reported having a positive experience building the selection sets for their activity: NT-NA-4: "*It was good. It was great. Like again, it was pretty self-explanatory. Yeah, it was pretty easy*" or, T-A-5: "*A really good experience actually. It was so intuitive, so nice...I felt like I was doing something that just felt good.*"

A few participants (3/24) provided rationale and explanations for the selection sets they created. One elaborated that his sets comprised of locations that have similar weather profiles or geographical profiles: NT-A-4: "*...because there were three sets to be made, I was actually thinking about dividing it into coastal, inland and...thinking about...hurricanes, cyclones...So maybe one [location set] would be just the East coast cities. One would be just the inland cities and one may be something else.*" The other two justified grouping locations that are geographically proximate to each other to increase the difficulty of the game: T-NA-4: "*I started choosing places that would be closer to each other because I figured that it would make it a little bit more challenging for...the players. Cause then you have...areas that are fairly close to each other and would have the same weather affect and so then it would make it a little bit more of a challenge. So that's where my selections were coming from.*"

And among the 12 participants who had access to the selection sets generator (SSG), 3/4 (9/12) actually used it. We looked whether the SSG affected the participants' sentiments regarding creating the selection sets. Of the 8 participants who reported a positive experience with Step 3, most of them (6/8) did not used the SSG. That was because 5 were in versions the ACW with no automation (T-NA, NT-NA) and 1 who had access to the SSG did not use it.

12.1.8. Scoring Formula Creation

In the fourth step of the ACW, participants had to create the rules that the system would use to evaluate and score the players' predictions. Creating the scoring formulas was the most difficult step for the majority (21/24) of participants. The biggest reason, two-third of them (14/21) reported, was that either they (4 responses) did not know what action to take: NT-NA-5: "*Even with instructions...I'm not sure what to do...I have no idea what...the sample formula was talking about...how do I replicate that? How do I know...what formula I should create.*" Or they (10 responses) were not sure whether the formulas created were appropriate, fair, or correct: NT-A-3: "*I just made up some formula, but I wasn't sure what I was doing. I mean writing the formula is easy, but what to write is difficult to think like...what formula best predicts the [outcome]...That was tough for me*" or, T-NA-4: "*I was trying to be like...if I did minus or plus this...would this actually come up to be a positive score and increase the score overall? So a lot of it was just me in my head trying to do [math]. I'm still not even sure if I'm accurate in what I ended eventually landing on. So yeah, that was complicated for me.*"

Another major reason this step was very difficult, according to 11 out of 21 respondents, was that it required background knowledge in mathematics or statistics (8 responses): T-A-1: "*the person just want to play the games. They don't want to know too* *much math problem. It's just a game for me, not the math problem*" or NT-A-3: "*I'm not good at mathematics and...statistical analysis.*" Or it required a background in the domain (3 responses): NT-A-1: "*if someone doesn't have...knowledge like in [climate]...they wouldn't know how to do it.*"

For 6 participants, creating the scoring formula was most complicated because of unfamiliar or unclear terms or names used either on the interface or in the explanations. Half of the responses (3/6) expressed that there was no context nor explanations for data properties (e.g. max temperature property in the weather dataset): NT-NA-5: "*I guess...I didn't know what a lot of these like variables (data properties e.g. max temperature) were or what they meant...I didn't know like how it would affect everything...there's just so much to like abs 100 minus abs. I've no idea what that is. No idea. So it's like [...] what is this doing? I don't know.*" Also, 2 of these participants opined the term 'dataobject' (e.g. ClimateDataObject) would be too complicated for average users: T-A-6: "*Like this 'data object' [as in 'ClimateDataObject' on the interface]...it sounds [like] programming.*" And one thought that the dataset and property names were too long to remember and type while creating the formula and proposed shortening them. For example HistoricalAverageDataObject_hightemperature could become HA_hightemp.

3 responses mentioned the user interface as another factor that complicated the creation of the formula. As shown in Figure 69, a participant tried to write her formula as an equation (by including the equal sign) which the system could not parse because it already assumed it. Another participant found the data property tree component unintuitive. And the last response elaborated on the disconnection between writing the

scoring formula and writing the formula explanation for the players; these actions were done on separate interfaces: NT-A-2: "*I didn't realize that the [formula] explanation box is below. I thought this interface...[the formula definition interface]. So [I] give the name [to the] defined the formula. Then I hit submit. It keeps saying I need to explain. I thought this [the formula itself] is the explanation. I guess the box on the previous page, the box at bottom saying explanation; somehow it's easy to miss that...because on this page [the formula definition interface] it feels like, Oh, it's done, but it's not. Like [I] got the error message twice and it's kind of confusing...So when I came here [the scoring formula step main interface] I thought, Oh, the job's done. But it's not...That's a little bit [of] disconnection I feel.*"

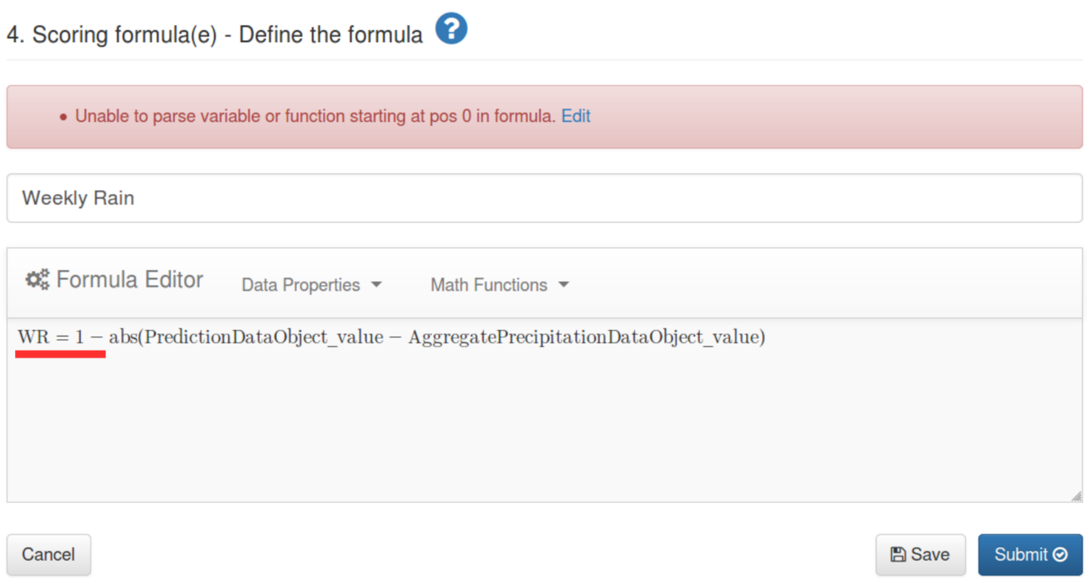

Figure 69. Formula Editor. A user is trying to write the scoring formula as an equation and the system rejecting it

One participant reported that a missing mathematical function from the ACW made writing his/her desired formula difficult: T-NA-3: "*I selected three* *parameters…[e.g.] maximum speed, minimum temperature. I did not really wanted to sum those. I wanted to compare with others then...select the minimum one, but I didn't see a function of minimum.*"

Feedback was given by 21 out of 24 participants. 81% (17/21) recommended that the ACW in step 4 (creating scoring formula) provide users with formula presets including their explanations and rationales: NT-A-1: "*If you had preset formulas for people and then had an option that you could make your own custom one. I think that'd be a lot easier,*" NT-A-6: "*Instead of you having to come up with the formula...maybe there's already like a built-in formula...maybe it can tell you...[the] name of the formula and then you click the button and then there's an explanation of what the formula does or like what it means...I feel as if that would have been a lot easier,*" or T-A-6: "*Maybe...give them...specific instructions on what can be [a] good formula...For people like that are not from statistics or math background...what can be some good formulas?... And instead of just saying there have to be two formulas...what are the variations that you can have?...So even though there was an example I feel...that it needs a little more but...not technical knowledge. So maybe more information on that can be given.*" A variation of the primary feedback suggested by another respondent was for the system to automatically generate the appropriate scoring formula based on certain criteria chosen by the activity author: T-A-1: "*I don't need to create the formula by myself. I just need to choose the criterion by myself...Show us the criterion in the system. And I can choose which criterion I...want to use.*" Five other participants suggested more detailed explanations of the data properties (e.g. max temperature, wind speed, humidity, etc...) like how they relate to or affect each other: NT-A-4: "*I wanted to use more data properties, but...I wasn't sure [about] how each of those data properties are linked...What one would need is...maybe a one or two line explanation...whatever that relation comes down to. So...if that is clear, maybe we can have more comprehensive formula. And that would be...something based on a scientific theory itself.*" The last suggestion from one participant was for the ACW formula editor to support symbolic representation of data properties in order to facilitate writing the formula: T-A-5: "*You...assign this max temperature A, minimum temperature B, and mean temperature X. Okay. So square root of A minus X, square roots of B minus X. It is very easy...if [people] can do square root of A - X, simple.*"

One participant provided the rationale behind his/her scoring formula. S/he explained that one of his/her goals was to reward players whose high temperature prediction scores and low temperature prediction scores were consistent: T-A-6: "*I want the person who's making that four and four unit difference [4 points for high temperature predictions and 4 points for low temperature prediction] to have more score or higher ranking than the one with seven and one unit difference [7 points for high temperature predictions and 1 point for low temperature prediction]...I was thinking...that difference [between the consistent people and non-consistent people] should be...taken into account in the scoring.*"

12.1.9. Communications Channels

In the sixth step of the activity authoring process, participants selected which communication channels (e.g. direct messages, instant messaging) will be available to the players during the activity. In general, the decisions made in this step by the participants were not the focus of their comments. Nevertheless, one participant explained that s/he selected only public channels for transparency and to prevent bullying or harassment. S/he elaborated: NT-A-5: "*I wasn't very sure about this direct messaging and instant messaging thing. So I just feel like that in all of these forums it would be good if people, whatever they are communicating is in...some kind of public forum or place. There isn't much direct messaging and stuff because I just feel like it does lead to other problems... I...just did not want any...negative sort of communication...Because the fact [is] when people feel like no one is seeing them or whatever they do, things can become a bit problematic...I guess with the whole issues of cyber bullying and stuff.*"

12.1.10. Created Prediction Activities

The table below describes features of the activities created by the participants using the ACW during Task 1. It focuses on the prediction schedule, the selection sets, and the scoring formulas. For the prediction schedule, the task asked for a prediction schedule with about 7 prediction rounds. The total average schedule size created by the participants was about 6 prediction rounds (mean = 6.37). For those whose used the automatic schedule builder the average was about 7 (mean = 7.40) and it was about 6 (mean = 6.10) for those who did not. Participants who created a very minimal prediction schedule (e.g. one round only) explained that they misunderstood the task instructions.

Table 5. The prediction activities created by the participants

Participant ID	Schedule (Prediction rounds)	Used ASB	Selection Sets	Used SSG	Used TPL	Scoring Formula	Notes
T-A-1	7	No	4	Yes	No	Incorrect	
T-A-2	2	No	6	Yes	No	Incorrect	
T-A-3	1	No	4	No	Yes	Incorrect	Only one prediction round because misunderstood task instructions. Did not change formula preset.
T-A-4	7	Yes	4	No	Yes	Incorrect	
T-A-5	7	No	3	Yes	Yes	Almost correct	Had the right idea but used wrong data property
T-A-6	6	Yes	8	Yes	Yes	Correct	Created a third formula for game balancing
NT-A-1	10	Yes	6	Yes	No	Correct	

Table 5. Continued.

Participant ID	**Schedule (Prediction rounds)**	**Used ASB**	**Selection Sets**	**Used SSG**	**Used TPL**	**Scoring Formula**	**Notes**
NT-A-2	6	Yes	4	Yes	No	Correct	
NT-A-3	7	No	6	Yes	No	Incorrect	
NT-A-4	8	No	3	No	No	Incorrect	
NT-A-5	8	Yes	6	Yes	No	Incorrect	Created 2 rounds using ASB but updated manually later
NT-A-6	7	No	5	Yes	No	Partially correct	Copied the example formula but did not have second formula
T-NA-6	1	No	5	No	Yes	Incorrect	Only one prediction round because misunderstood task instructions. Did not change formula preset.

Table 5. Continued.

Participant ID	Schedule (Prediction rounds)	Used ASB	Selection Sets	Used SSG	Used TPL	Scoring Formula	Notes
T-NA-1	8	No	3	No	No	Correct	Inspired by the example formula
T-NA-2	1	No	3	No	Yes	Correct	Only one prediction round because misunderstood task instructions. Inspired by the example formula
T-NA-3	13	No	4	No	Yes	Incorrect	Did not change the selection sets presets. Did not change formula preset.
T-NA-4	8	No	7	No	No	Correct	Reproduced the formulas from the Fantasy Climate demo

Table 5. Continued.

Participant ID	Schedule (Prediction rounds)	Used ASB	Selection Sets	Used SSG	Used TPL	Scoring Formula	Notes
T-NA-5	7	No	4	No	Yes	Almost correct	Had the right idea but used wrong data property. Did not change the selection sets presets
NT-NA-1	7	No	3	No	No	Incorrect	
NT-NA-4	8	No	3	No	No	Correct	
NT-NA-2	9	No	3	No	No	Incorrect	
NT-NA-3	1	No	3	No	No	Correct	Only one prediction round because misunderstood task instructions
NT-NA-5	7	No	3	No	No	Correct	Inspired by the example formula
NT-NA-6	7	No	3	No	No	Correct	Inspired by the example formula

Regarding selection sets, there was no specified requirement for the number or size of the selection sets. On average, the participants created about 4 selection sets (mean = 4.24). The average number of sets for the participants who used automation was a bit above 5 (mean = 5.33) and a bit below 4 (mean = 3.67) for those who did not. We also examined the created selection sets of the 11 participants who used no automated tools or manually edited the template's preset. While 4 of them followed the example selection set for Fantasy Climate where the locations were spread out away from each other, the other 7 custom sets deviated from it. For example, NT-NA-6 created 3 custom sets comprising of eastern coastal cities, southern cities, non-coastal northern cities as shown in Figure 70.

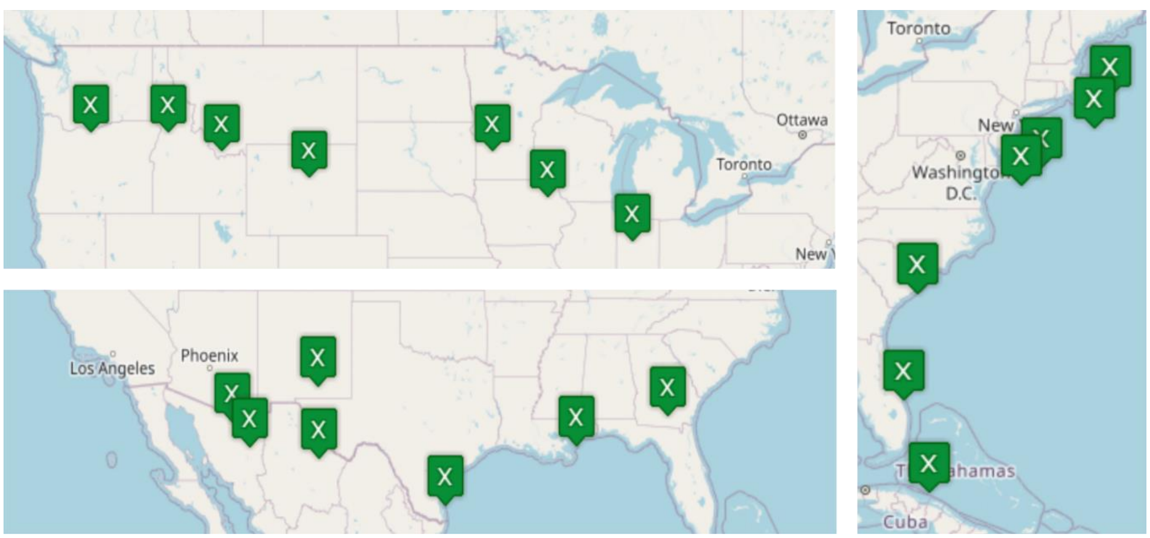

Figure 70. Custom selection sets created by a participant

For the scoring formula, the task specified that the formulas for the prediction activity should reward players whose high/low temperature predictions deviate the most from the historical average high/low temperature. We examined the scoring formulas that the participants created labeled them 'Correct' or 'Incorrect' depending on whether they met the task specification. Ten of the 24 formulas were correct including one participant who reproduced the formulas from the Fantasy Climate demonstration and three who were inspired by the form the example formula in the explanations. Two participants had the correct idea but selected an inappropriate property for their formula. Both used the HistoricalAverageDataObject_meantemperature attribute, which is the average temperature for the day in the historical record rather than the average high temperature for the day. One participant copied the example formula but did not produce the second required formula. In the 11 out of 24 formulas that were incorrect, 2 were the unchanged template preset.

12.1.11. ACW Limitations

Were there aspects of the prediction activity participants wished to customize but such customizations were not possible or limited with the current version of the ACW? 20 of the 24 participants answered that there was nothing they wanted to do during this task but was not possible. The suggestions of the remaining four were: (1) a customizable reward system to motivate the players in the game, (2) customizing the aesthetics of the prediction game interfaces, (3) customizing the name of the created selection sets (since the name is currently autogenerated e.g. 'Selection Set #2' – see Figure 68) and, (4) adding a description for selection sets elements (in the case Fantasy Climate, locations): NT-NA-1: “*maybe we can add descriptions of the different cities that I could choose from. How, like if there's a different page...that pops up where you can read about...the climate...that the cities had, before you could choose a city.*”

12.2. Discussion

The responses from participants provide insight into the role of background knowledge in the authoring process, the effects of the template and automation in ACW, and the difficulty and order of steps when authoring an activity.

12.2.1. Was Background Knowledge Required to Use the ACW to Author an Activity?

We computed the correlation between the participants background (educational, knowledge of statistics, gaming and, fantasy sports) as reported on the pre-questionnaire and the results reported in the previous section. We found none. There was no correlation between background and task duration. This confirms the participants' responses that no background in gaming or fantasy sports was necessary to use the ACW. We were particularly concerned about assumed knowledge of fantasy sports. But, given that prediction games were modeled after fantasy sports and half of our participants had no fantasy sports experience, there does not appear to be a strong need for such experience. One participant mentioned that a background and interest in the domain (climate/weather) may help. This makes sense as such knowledge is valuable for understanding the domain data and for deciding on a specific prediction task. Perhaps weather and climate, at least at the level of observed temperatures, are well enough understood that the common nature of this domain, masked any further need for domain expertise.

Some participants posited that the scoring formula step was difficult because it required a background in mathematics. Nevertheless, there were no correlation between the education background (degree and field) or knowledge of statistics and the correctness of the formulas created by the participants. Creating the scoring formula clearly requires an understanding of basic math (addition, subtraction, average/mean, etc.) but much of the perceived complexity appeared related to authoring the formula in a structure editor, which is most often found in programming environments.

Overall, we surmise little background knowledge is required to use the ACW to author a prediction activity. Based on the comments above, the structure provided by the ACW combined with the automation tools, templates, and explanations including context, tutorials and examples help bridge the knowledge gap and guide the users during the authoring process. The following comments illustrate: NT-A-1: "*I felt like it [the ACW] was...basically holding your hand through the whole thing*", T-A-5: "*[I think] anybody can do this. Anytime they can, they can do this*", and T-NA-5: "*you don't need any background knowledge to do this type of stuff.*"

12.2.2. What Were the Effects of the Template and the Automated Tools?

Participants were not required to use the template or the automation tools during the experiment. As a consequence, even when available, some still chose to manually customize the prediction game for different reasons such as misinterpretation of instructions, visibility or control. However, for many their experience were or would have been enhanced with the template or automation

Figure 61 shows that the template helped to reduce the overall task duration. Even though they were divided on whether the template component was indispensable, nearly all the participants agreed that it was helpful (Figure 54). T-NA-2 said: "*The* *template was really helpful...It basically, it had that formula in there...That really guided me.*"

Similarly Figure 61 also shows that less time was spent in Step 2 (building the prediction schedule) and in Step 3 (creating the selection sets) when the automation tools, the ASB and the SSG, were employed respectively. One suggestion for the prediction schedule step proposed by the participants who did not have access to the ASB was for a tool similar to it. However, while only one participant who reported a positive experience in Step 2 used the ASB, nearly all who expressed negative comments did not use the ASB. Likewise, most of the participants who made positive comments on their experience in Step 3 did not use the SSG. This suggests that although the automation tools may not be very important to the authoring process but they enhance the user experience as the following participants commented: NT-A-3: "*I would have gone for it [the automatic schedule builder] if I saw it more carefully and understood how it worked...This is easier*" and T-A-6: "*I didn't use the manual one. Probably that would have been a little more time consuming, but with this [the selection sets generator, it] was quicker*."

Table 4 shows that among the participants who reported a positive experience with authoring the ACW, 60% of them used either the template or automation or both. The rest (40%) used neither, either because they were in versions of the ACW with no template nor automation or for reasons already discussed in sections 12.1.4 and 12.1.6. On the other hand, half of the four participants with a negative sentiment about their experience did not have access to the template or automation.

Thus, even though the template and the automated tools enhanced the participants' experience, they were not indispensable during the authoring process.

12.2.3. How Did the Explanation System Affect the Authoring Task?

For any new user of the ACW, an explanation dialog automatically popped up at the beginning of every step. Participants' opinions were divided: half thought the popup dialog was bothersome and the others thought it helped at least in some ways. This corroborates the results of the questionnaire (Figure 57). We posit that the division may be due to personal preferences. Some people may need more hand holding than others. One way to overcome this division as some participants suggested is to create an explanation panel that is constant on every interface of the ACW as shown in Figure 63. This way even for new users, explanations will be shown on every interface anytime without blocking it.

However, confirming the results of questionnaire (Figure 57) once more, a majority of participants affirmed that the explanations helped them accomplish their task. Or in the case of the scoring formula step, more explanations would have helped. The main recommendation for this step was the availability of more example formulas accompanied by their explanations: descriptions, justifications and impacts on the game.

At the same time, half of the participants also commented that explanations were wordy. In future iterations of the ACW, we will need to revise the explanations content to find a balance between concise information and important information. We may also deepen the way we progressively disclose information to the users. And given the volume of information in the explanations content, the system may feature a search functionality to help the users reach the desired information quickly or easily.

Even though most agreed that the explanations were easy to understand (Figure 57), one participant reported unfamiliar or unclear language in the explanations content that is worth considering. This theme of unfamiliar, confusing, complex terminology in the ACW recurred in many responses especially when describing the experience building the prediction schedule or creating the scoring formula. The reason is that the explanations content or terms on the user interface were produced by the system developer where his technical background likely influenced his choice of vocabulary. For example, participant T-NA-4 reported struggling with the term 'persist' which according to her/him: "*persistence is struggling or persevering through something*" but the developer meant 'persist' as in storing information (e.g. in a database or on a hard drive). Or participant T-A-6 pointing out that the term 'dataobject' sounds like programming. Perhaps future revisions of the explanation contents and the user interfaces of the ACW should be done in collaboration with people with little to no technical background in software development. Thus, as we expected, the explanations aided the participants authoring of their prediction activity using the ACW. However, they need improvements in user interface, content quality (including terminology) and content quantity.

12.2.4. Were There Alternatives to the Current Step Order of the ACW?

The sequence of the activity authoring process was logical to almost every participant. However, a few responses revealed alternative ways of thinking about the

ACW. T-NA-4 said: "*I could see doing the selections before scheduling and then having the location [sets] built into the scheduling aspect. So then saying like on this week you're going to go with location set one, on this [other] week you gonna go with location set two, and that way you can finagle it because on this level you kind of don't have any control of what location set goes with what schedule.*" This comment is interesting because it challenges our current design of the ACW based on the idea that the authoring process of a prediction activity is a sequence of discrete steps. The stepwise process may undermine the actual need to co-design aspects of the activity authored in separate steps.

Participant T-NA-4 does not only want to do Step 3 (selection sets) before Step 2 (prediction schedule) but provides a scenario in which prediction dates are coupled with selection sets. In other words, the steps of the authoring process are not distinct and separate from each other. Instead, there are many ways, they can be interdependent. Another participant's comment shows another example: T-NA-1 said: "*I tried [creating two] selection sets. I thought one would be for the highest [temperature prediction and one would be for the lowest temperature prediction]...I wasn't sure the same set can be used [for both type of predictions]...It wasn't written anywhere...[that] this set would be used for the highest and the lowest [prediction]."* The study task asked the participants to create two scoring formulas for their Fantasy Climate activity: one for high temperatures and another for low temperatures. Participant T-NA-1 only created two selection sets (ACW Step 3) because in her/his mind, selection sets mapped to prediction formulas (ACW Step 4). Synthesizing the examples of participant T-NA-4 and participant T-NA-1, we can imagine a winter prediction activity where players are asked to predict:

- in one week, from a set of northern cities, the location with the highest weekly precipitation (snowfall) and the location with the lowest weekly precipitation,
- and in the following week, from a set of coastal cities, the location whose high temperature deviates most from historical average and the location whose high temperature deviates the least from historical average.

Authoring such an activity requires that the prediction schedule (ACW Step 2), the selection sets (ACW Step 3), and the scoring formulas (ACW Step 4) be interdependent, which the current version of the ACW does not support as participant T-NA-4 has pointed out. Future iterations of the ACW will have to consider alternative ways of designing prediction activities.

Another comment challenged our perspective of the prediction activity authoring process as a solitary task. The participant was explaining why creating the prediction schedule was very difficult for him/her: T-A-1: "*I should also choose a person I want to compete with and if I chose a date by myself...I need to communicate with that person. The person I choose maybe, I don't know. So I don't want to talk...more with the person...about the exact dates: the due dates and the score dates.*" In an informal setting where a group of people want to create a prediction activity to play with each other, many aspects of the design may need to be negotiated such as the prediction schedule, which locations should be part of the selection sets in the case of Fantasy Climate, the complexity of the scoring formula, which analysis tools shall be available during the game, etc... In other words, authoring a prediction activity may be a collaborative task requiring cooperation among the stakeholders. As a consequence, future iterations of the ACW may begin with setting up the activity players (ACW Step 7) instead of ending with it. T-NA-4 said: "*I mean I could see the social aspect, [setting up activity players,] being chosen at the beginning before you're actually [customizing] everything since that's just kind of determining...who you're adding to this beforehand.*"

As indicated by the feedback, participants envisioned alternative ways of authoring a prediction activity that challenged the current design of the ACW where (1) some steps are interdependent, and (2) authoring becomes a collaborative activity.

12.2.5. What Were the Most Difficult Steps and Why?

The prediction schedule, the selection sets and the scoring formulas are the most consequential components of a prediction activity. Unsurprisingly, Step 2 (making the prediction schedule), Step 3 (building the selection sets) and Step 4 (creating the scoring formulas) were the most difficult steps for the participants during the authoring process. This was reflected by the participants' time spent in these steps (Figure 60).

Making the prediction schedule was found tedious by many participants. As a result, the main suggestion was for a tool similar to the automatic schedule builder. Manually creating prediction due dates coupled with scoring dates for an activity with a long duration (e.g. a semester) would be (a) repetitive, and (b) difficult to keep the dates well organized. In the next iteration of the prediction schedule step of the ACW, the automatic schedule builder should be the default user interface. For authors who desire more control, the user could manually edit the schedule generated by the builder or explicitly choose to not use the builder and author the schedule without its support.

Among the three difficult ACW steps the participants discussed, the third step (creating the selection sets) seemed the least difficult. Besides few user interface issues, the participants mainly complained about undocumented system requirements. The average time spent on this step was the lowest, and only one participant disagreed that creating the selection sets was easy according to the post questionnaire (Figure 59). However, review of the activities created indicates that many participants did not understand the purpose of this step or how the customizations therein would affect the game. There were few mentions of the role of selection sets in participants' responses and they did not seem to reflect on them. The vast majority of the participants who had access to the Selection Set Generation, or SSG, (in versions T-A and NT-A) used it (which would explain the lower time spent compare to making the prediction schedule) and made no change to the selection sets that it automatically generated. Among the participants who used the template, half left the selection sets presets intact and most who updated the presets only made minor cosmetic changes. NT-A-2 said when discussing her experience with creating the selection sets: "*First of all, what am I supposed to do? Kind of scary to me. I just went for the auto generator. That's easy and helpful*", and T-NA-3 who used the template presets: "*I'm from India, I don't know which city in the US are the...coolest...if I had a fair idea of what cities can be cool...I would have [created] my own custom sets.*" Future support for creating selection sets should be more transparent by including explanations of sets in templates and access to information about potential elements of the pivot set (e.g. set of locations in Fantasy Climate.) These could include descriptions and examples of effects on a prediction activity as NT-NA-1 suggested: “*Maybe we can add descriptions of the different cities that I could choose from...like if there's a different page...that pops up where you can read about...the climate...that the cities had, before you could choose a city.*” Additionally, expanding the SSG to enable authors to specify additional features to consider when generating selection sets would allow it to provide support across a broader range of activity goals. For example, the SSG could generate sets based on requested similarities or differences in elements’ geographical region (e.g. ‘southwest’ or a set of geocoordinates), temperature range, wind speed range, or even frequency of extreme weather events.

Creating the scoring formulas was the most challenging step for nearly all participants. This step involved at least some basic mathematics and statistics as about half of the participants confirmed. However, the role of math is probably not the reason this step was the most difficult. Only one participant reported having no experience with statistics (Figure 53), and a majority of participants was either in an engineering or science field signifying a basic knowledge of mathematics. We surmised that this step was very challenging because the effects of the scoring formulas on the prediction activity were very opaque, and there was not enough information to help the participants understand the consequences of their customization on the game.

Did participants understand the impact of their decisions? The prediction schedule effects were clear and obvious to the participants. In the case of the selection sets, the participants with access could use the SSG if the consequences of their choices were obscure to them. However, the question of how the scoring formula affected the prediction activity was very difficult to answer and the template presets and the explanations were insufficient. Some participants reported that, even with explanations, they did not know what formula to write. And for those who had an idea of what formula to write, there were still questions that were very difficult to answer. For instance, are the scoring formulas correct? In other words, do they reward good predictions accordingly? NT-NA-4: "*I don't know if I did that one correct or not...I was like, I am not sure what I am supposed to put here...I think like I just wasn't 100% sure. Like how...I was supposed to...make sure that the scores...would be higher...if they (the players) got it correct. So I'm just kind of hoping that I was right.*" Or will the formula create a fair and balanced (ultimately fun) game and not over-reward certain desirable player behaviors over others? NT-NA-1: "*so I didn't know what to do with the formula and how to set it up so that it would be fair for the players.*" Unsurprisingly many participants' feedback went beyond the suggestion of more presets formulas to choose from but suggested accompanying explanations as to why such formulas are 'good' or appropriate, and how they affect the prediction activity. Future work should build on this feedback and explore methods to simulate and visualize the effects the scoring formulas. One participant during a pilot study explained that she wished she could simulate the formulas she had created, and update them based the results. She would repeat the process until she obtained the formulas that satisfied her.

Another reason creating the scoring formulas was reported as the most difficult might had been because it was not engaging. At least several participants reported having a positive experience with making the prediction schedule and creating the selection sets. No one made a positive comment about the scoring formula step of the ACW. We guess that what many participants meant by this step requiring a mathematics background was that creating the scoring formula felt like mathematics homework. As participant T-A-1 insisted that it was just a game for him/her and "*not the math problem*" or NT-A-2: "*...some of the functions might look daunting to them [other users]. Wow! Doing math again.*" As many suggested, one way to circumvent this hurdle in the next iteration of the ACW is to provide example formulas with more detailed explanations that users can select among. Further out, it might be possible to automatically generate an initial scoring formula based on criteria configured by the users (e.g. which data features to include) combined with an analysis of the historical data.

From participant responses, it appears that the most difficult steps of the ACW were Step 2 (making the prediction schedule) mainly because it was tedious, Step 3 (building the selection sets) because its implications were a little bit obscure but could be circumvented with the SSG, and Step 4 (creating the scoring formula) which was the most challenging because its effects on the prediction activity were opaque and there were insufficient support to overcome it.

12.2.6. Usability Issues

Besides confusing or complex terminology, issues with the ACW user interface also appeared as a theme in the participants' description of their experience authoring the prediction activity. Although the issues raised were unique to each step, it would behoove us to incorporate the participants' suggestions in the future improvements of the ACW user interfaces. For instance, we may move the formula explanation textbox next to the formula editor in the formula definition interface for cohesiveness, update signifiers, icons, and labels in the selection sets interface to reduce confusion, update the prediction schedule interface to set time only once so that schedule making focuses on creating dates only, and make system requirements and their rationales (e.g. the minimum size of a selection set must be 5) more transparent.

13. CREATING PREDICTION GAMES: INTERVIEW RESULTS

In the second part of the study, the participants designed and described a prediction game in the domain of their liking (task description in Appendix G). We interviewed them afterward to find out more about their game (interview questions in Appendix K). The goals were to discover what domains and prediction activities people find engaging, and why. Additionally, the domains and designs enable reflection on how the ACW would need to be expanded to author such activities.

For our analysis, we examined the 24 games produced by the participants and categorized them. The games could be grouped into six game categories: pop culture, politics, commodity/stock prices, education, sports, consumer products. An additional category, N/A, was assigned for two games that were not prediction games and did not meet the specifications of the assigned task. However, the motivations behind these games are worth considering.

Figure 71 illustrates the categories extracted from the 24 games produced by the participants. The most common domain (9/24) was pop culture and media which included television programming, movies, music, gaming, etc... Many games in this domain were about predicting the outcomes of a competition television show like American Idol. In the four political games, players had to either predict the winner of an election or the voting patterns of the election. Four participants also made games about predicting commodity or stock prices. The education domain consisted of games where players had to predict either students' overall performance (GPA) or their performance on an evaluation (e.g. test or assignment). There was in a single game in sports (wrestling), and another single game in consumer products (predicting sales.) Two participants created games that were not prediction games (N/A). For example, one of them was to compete with peers to (physically) clean plastic bottles off Antarctica. The following table lists the prediction activities with their brief description.

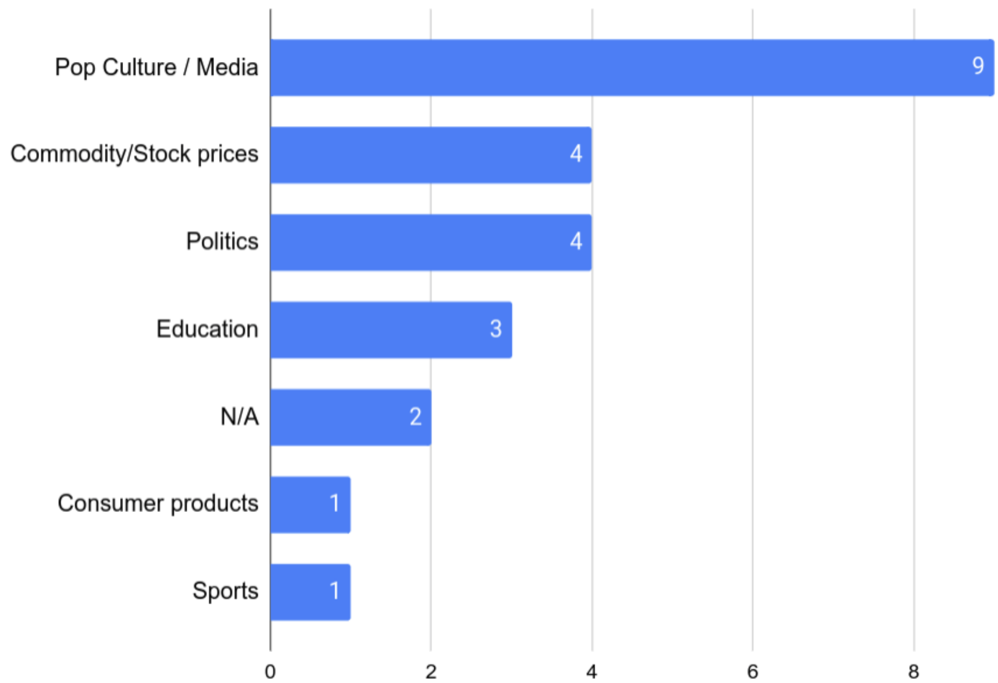

Figure 71. Domain of games created by the participants

Table 6. List of games created by the participants

Game	Category	Participant	Brief Description
G01	Consumer products	T-A-1	Predict the volume (sold) and, the rank among other smartphone brands of IPhone new model on its release date

Table 6. Continued.

Game	Category	Participant	Brief Description
G02	N/A	T-A-2	Compete with peers to (physically) clean plastic bottles off Antarctica
G03	Commodity/Stock prices	T-A-4	Predict the price of gold
G04	Pop culture / Media	T-A-5	Predict the box office of a new Marvel movie for its release weekend
G05	Pop culture / Media	NT-A-1	Predict the rank of artists on the Billboard hits weekly
G06	Pop culture / Media	NT-A-3	Predict weekly who will be eliminated on the television show The Voice
G07	Politics	NT-A-6	Predict the poll number (percentage) of a political candidate during a race
G08	Education	T-NA-3	Predict a student grade weekly (or after any assessment e.g. quizzes, homework, etc...) throughout the semester
G09	Commodity/Stock prices	T-NA-4	Predict the stock price of popular companies e.g. Amazon
G10	Education	T-NA-5	Predict the rank of students in a reading competition
G11	Education	NT-NA-1	Rank students by their GPA biweekly
G12	Pop culture / Media	NT-NA-4	Predict weekly who will be eliminated on the television show The Bachelor

Table 6. Continued.

Game	Category	Participant	Brief Description
G13	Pop culture / Media	NT-NA-2	Predict weekly the rank of contestants on American Idol and, the contestant who will be eliminated
G14	Politics	NT-NA-3	Predict the result of an election (the percentage of votes) obtained by each candidate or political party)
G15	Pop culture / Media	T-A-3	Predict the best actor over a period of time
G16	Commodity/Stock prices	T-A-6	Predict the oil price in a given country on a given date
G17	Politics	NT-A-2	Predict who will win the 2020 presidential election
G18	N/A	NT-A-4	Predict in realtime what your partner will say in an Instant Messaging conversation
G19	Pop culture / Media	NT-A-5	Predict who will win a competition show like American Idol
G20	Pop culture / Media	T-NA-1	Predict the team who will win a PUBG game
G21	Politics	T-NA-2	Predict who will win the 2020 presidential election
G22	Sports	T-NA-6	Predict the winner of a WWE Smackdown match
G23	Pop culture / Media	NT-NA-5	Select the team of contestants to be successful in American Idol or the Voice

Table 6. Continued.

Game	Category	Participant	Brief Description
G24	Commodity/Stock prices	NT-NA-6	Predict the average airline ticket price to different given locations from a given location

13.1. Information Resources

Participants were asked about the sources of information or data that the players of their game would likely use to make predictions. Some participants referred to more than one potential source of information or data. The single N/A in Figure 72 refers to the participant who misunderstood the idea of a prediction game.

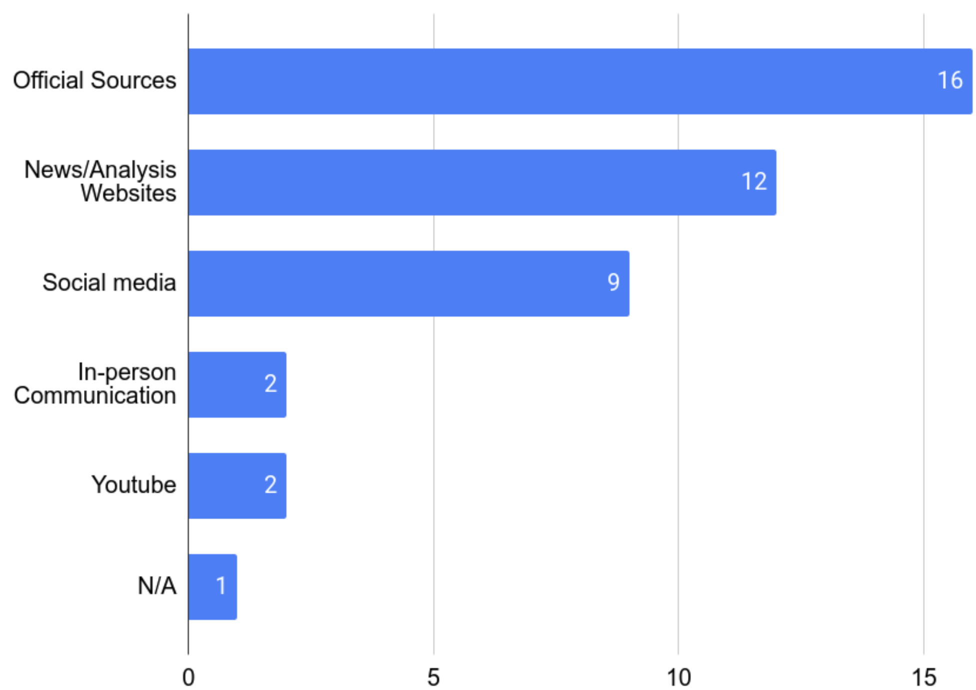

Figure 72. Information resources of the created prediction games

Almost 2/3 of participants (16/24) reported that the players of their game would use official sources for prediction making. The sources are for example the official website of political candidate (G07, G14), the television show itself or its official website (G06, G12), or the official website of a corporation (G09).

For half of the games (12/24), participants affirmed that news or analysis websites would be beneficial to the players during prediction making. According to T-A-5 who created G04, the Marvel movie box office prediction game: "*...news media outlets will go create hype about the movie. So that is another source of the other of data...for the movie.*" Participant T-NA-4 also highlighted as a source of information for his/her game (G09, the stock prices prediction game): Google Finance (https://finance.google.com/finance), a stock analysis website that includes visualizations of stock price trends.

Another source of information provided by 9 participants for their game is social media sites such as Facebook or Twitter. For instance, NT-A-1 the creator of G05, the Billboard music ranking game said: "*...social media is a good way to know...so things to take into account...if like someone is...doing something really controversial, people are going to listen to their music because they're...in the news...*"

Two participants suggested that in-person communication could be a source of information for their game. T-A-4 the creator of G03, a game about predicting the price of gold, explained: "*...if somebody...part of that group meeting...like potluck...really need it, they come back to the group and ask, 'Hey guys, do you really know what the analysis of the gold? Is really the minister or someone doing something [in] the news?'*" For his/her game (G04, the Marvel movie box office prediction game), T-A-5 posited: "*...generally people are smarter [than] the news. They talk to their friend...anticipation that their friends are also Marvel movie fan. So the anticipation of the fans then there's a hype around the movie.*"

Youtube (youtube.com) was mentioned by two participants as a source of information for their game. For G01, the game about predicting the number of iPhones sold, T-A-1 commented: "*YouTube, for example, Monday, people will provide some videos about the new iPhone*" and for G13, the American Idol contestants ranking game, NT-NA-2 said: "*News from those [social media] and all the fans commenting in the Youtube. They [the show's production company] will have the videos uploaded or something like that.*"

13.2. Data Analysis Tools

9 participants proposed 2 kinds of data analysis tools to aid the players of their game during prediction making. One of them was a comparison tool proposed by 2 participants. For G12, a prediction game based on the outcome of reality show, NT-NA-4 elaborated: "*...maybe there could just be, a compare, a chart of like all the girls. And so you could...compare which girls like have already left and...which girls have stayed...So you could see...this is kind of...the type of girl that he's...tending to keep, and...not to keep.*" The other type was a visualization tool proposed by 7 participants. For example, a chart illustrating the popularity of candidates by visualizing their poll number (G07) or Google Finance that features visualizations of stock prices over time (G09).

13.3. Prediction Modes

A prediction mode is the method through which players of a prediction game input their predictions. 10 out of 24 participants created games where players make predictions by entering a numerical value. For example in G03 or in G09, the players will play by inputting the price of gold or, stocks respectively. In 8 games, players predicted by selecting one or more items from a given set. For example in G12, the players will select the contestants who will be eliminated from the competition. And in 5 prediction games, the predictions are made by ranking items from a defined set according to a given metric. For example in G11 or G05, the players would rank students by their GPA or, rank artists by number of downloads respectively.

13.4. Data

Prediction games were originally conceptualized as a straightforward mapping from the model of fantasy sports to other data-intensive domains. Thus, assumptions about the characteristics of prediction games were often features of fantasy sports. In particular, it was assumed that there would be an external provider of historical and upcoming data, that players would be scored based on how well their prediction matched actual observations, and that there would be competition among players.

Table 7. Data properties of the games created by the participants

Participant ID	Data producer	Historical data identified	Data source for upcoming event identified	Scoring makes use of real event data and prediction	Includes competition among players
T-A-1	External	Yes	Yes	Yes	Yes

Table 7. Continued.

Participant ID	Data producer	Historical data identified	Data source for upcoming event identified	Scoring makes use of real event data and prediction	Includes competition among players
T-A-2	?	?	?	?	Yes
T-A-3	External	Yes	No	No	Yes
T-A-4	External	Yes	Yes	Yes	Yes
T-A-5	External	Yes	Yes	Yes (F)	Yes
T-A-6	External	Yes	Yes	Yes (F)	Yes
NT-A-1	External	Yes	Yes	Yes	Yes
NT-A-2	External	Yes	Yes - Unstructured	Yes	Yes
NT-A-3	External	No	Yes - Unstructured	Yes	Yes
NT-A-4	Personal	Yes	?	?	No
NT-A-5	External	No	Yes - Unstructured	Yes	Yes
NT-A-6	External	Yes	Yes - Unstructured	Yes	Yes
T-NA-6	External	Yes	Yes	Yes	Yes
T-NA-1	External	Yes	Yes	Yes (F)	Yes
T-NA-2	External	Yes	Yes - Unstructured	Yes	Yes
T-NA-3	Personal	No	Yes	Yes (F)	No

Table 7. Continued.

Participant ID	Data producer	Historical data identified	Data source for upcoming event identified	Scoring makes use of real event data and prediction	Includes competition among players
T-NA-4	External	Yes	Yes	Yes (F)	Yes
T-NA-5	Personal	No	Yes	Yes	Yes
NT-NA-1	External	Yes	Yes	Yes	Yes
NT-NA-4	External	No	Yes - Unstructured	Yes	Yes
NT-NA-2	External	No	Yes - Unstructured	Yes	Yes
NT-NA-3	External	Yes	Yes	Yes (F)	Yes
NT-NA-5	External	No	Yes - Unstructured	Yes	Yes
NT-NA-6	External	Yes	Yes	Yes	Yes

The table above shows these characteristics of the games created by the participants. For 80% (20/24) of the games, the data was produced by an external source like a website recording and displaying the price of gold (G03) over time or the weekly rank of top artists (G05). Whereas for the rest, the players themselves would produce the data. For example, G08 is based on the grades of the student who is playing the game. Next, about a third (7/24) of the games created had no historical data available. For example, for the games based on competition television shows (G19, G06: who will win or be eliminated from The Voice or American Idol or, G12: who will be eliminated from The Bachelor), there is barely historical information on the contestants to be examined by the players like the number of competitions they have won in the past, the number of years they have been singing, etc... Even if there is a little history available on some contestants, it may be insufficient to inform good predictions. There were historical data available for the rest (16/24). Almost all (21/24) the games featured data for the upcoming event used for scoring like the results of an election. The upcoming data was unstructured for eight of these. Once more the competition shows games are good examples. The source of the upcoming data is the television show itself and the official website of the show. However, such data is not organized. Only G15, the game about predicting the best actor over a period of time, did not feature data for the upcoming event since the scoring system evaluates predictions on what an actor has done. Nevertheless, the scoring system of nearly all (21/24) the games made use of the player's prediction and the data for the upcoming event. For six of these, the participants even formulated the scoring rules: NT-NA-3 (G14, the game about predicting the result of an election): *100 – Abs(predicted vote percentage – actual vote percentage) for every candidate*. And only G08, the game about predicting one's grade, did not include competition.

13.5. Purpose / Context

We also discussed with participants the context into which their game would be created or played. Some participants had more than one purpose for the game they designed. Our examination showed that 13 participants viewed their game as an extension of the culture like fantasy sports is an extension of sports. For example, responding the question about the audience of their game, NT-A-1 (G05, the Billboard music ranking game) said “*people who like music the younger generation is more involved in, you know, social media and, you know, stuff like that*” and NT-NA-2 (G13, the American Idol contestants ranking game) corroborated: “*Whoever is a fan of, or maybe who likes...music. Mainly [may be young adults who] enjoys watching TV shows.*” And T-A-5 described the motivation behind his prediction game (G04, the Marvel movie box office prediction game): “*So Marvel brings out movies and their movies are really good. And people are dedicated to those movies, like...Avengers. So they earn a lot of money...because they are very popular.*”

9 participants suggested that their game could be a part of their social activities: NT-A-6 (G07, the game about predicting the poll numbers of a candidate): “*this [prediction game idea] is something that came into my mind because I am in a political organization here on campus...we try to find out about different stuff going on today nowadays...the example...reminded me of it. And with my friends, we're like, Oh my God, who's going to get the nomination...the Republican side...several friends...are interested in this*” or, NT-NA-4 (G12, the game about predicting the outcome of the show The Bachelor): “*I think it'd be fun for like college girls. So we could like play altogether.*”

7 participants explained that their game may have educational benefits by either helping people learn more about the game's domain (e.g. politics): NT-A-6 (G07, the game about predicting the poll numbers of a candidate): “*A bunch of much of law students. Just kidding. A bunch of people...are really interested in...what's going on in the political climate, because for some people it's like a game...some [other] people actually just...want to know what's happening*” or motivating people to learn in general:

T-NA-5 (G10, the game about ranking the contestants of a reading competition): "*So the game I created, the goal is so encourage student learning - that's why I started with education because I think this type of stuff, it can be a good motivation for education.*"

Finally, T-A-1 expressed that one of the main interests of her game (G01, the game about predicting the number of iPhones sold) was commercial or marketing purposes: "*Apple will sponsor this game to make people more interested in their brand.*" For T-NA-3, his/her game (G08, the game about weekly predicting a student grade) was a practical tool to help students be aware their grade throughout the semester: "*So in this game, basically, a student will be able to...track how much he needs to score in the upcoming tests in order to achieve an A.*"

13.6. Discussion

The results above indicate the prediction games that the participants found engaging need not be in a domain where historical data is available. The activities based on reality television or competition game shows are an interesting case. There is virtually no historical data available. As a consequence, the players will likely rely more on information resources such as news to make predictions. Furthermore, the data (e.g. set of contestants) driving these games is unstructured. Designing such games will require creating formal structure of the data that can be programmed into the system. Another aspect of these games is that the pivot set changes for every prediction activity because there are new contestants every season. This poses new challenges in authoring activities for such games as the new pivot sets must be ingested into the system for every authoring task. Another result is the type of games where the data driving them was produced by the players about themselves as T-NA-3, the author of G08, commented that the game could be use a practical tool to keep track of one's grades. Thus, our perception of prediction games is broadened in that they need not be driven by structured data nor have historical data and may be driven data produced by the players.

These results also provide some insights into what make prediction games engaging. Nearly all the games we examined were grounded in socio-cultural contexts that the participants and their friends found engaging or entertaining. T-A-4's game (G03) about predicting the price of gold is a very good example. Gold plays an important role in her culture: "*But for us, when the wedding season or some festival season comes, the demand [for gold] will be really high. People don't want to buy, but still buy as like customs, rituals...And sometime we [are] getting a new baby...you need to buy*." As a consequence, it is a main topic of conversation during socialization: "*It's like this when the community group is happening...people keep want predicting on every talk.*" S/he provided an example: "*It's like potluck or something, then somebody...really need it, they come back to the group and ask, 'Hey guys, do you really know what the analysis of the gold?' Is really the minister or someone doing something [in] the news.*" In fact, the participant explained that she has already been playing a version of her prediction game with her peers and family: "*We...[as] a group [in WhatsApp]...talk about that game...we have that idea...let's make this game very popular.*" Thus, his/her prediction game was not only based on a popular aspect of his/her culture but a formalization of a social activity she already engages in with her family and peers. This is essentially true with fantasy sports. Fantasy sports are engaging because sports in general are entertaining and popular and, several studies including ours have shown that socializing was a primary factor that motivated people to play the game. Hence, grounding prediction games in popular socio-cultural contexts (e.g. G04, G06, G12, G13) may make them engaging.

For domains that do not have the popular or entertainment appeal (e.g. climate, politics), designers may leverage the social aspect of prediction games to make them more engaging. They may do so not only through competition or, communication tools that afford social interaction during the game like in fantasy sports, but also through how and where players seek information to make prediction. In about half of the games, peers were a source of information for prediction making whether in person (e.g. G04) or through social media (G05). Thus, a future feature of the prediction games system may be the ability to stream and display news and information from major social media sites like Twitter or Facebook.

Finally, another lesson from the results in the previous section is the recognition by some participants of the educational benefits of prediction games. One of the education benefits is players learning more about the game's domain. When asked who was the audience of his/her game (G07, the game about predicting the poll numbers of a candidate), NT-A-6 replied immediately: "*A bunch...of law students.*" S/he went to argue that although politics was a game for some people, others really wanted to learn about it. T-NA-4 gave a similar response (G09, the stock prices prediction game): "*...business majors...You also might just get people interested in...investing.*"

Another educational benefit is players learning more about a different domain that is related or affects the game's domain. In G03, the gold price prediction game, the players will not learn about gold itself (e.g. chemical properties, how it is mined or how it is priced). But instead like T-A-4 and his/her peers, they will learn more about global socio-political events since they affect the price of gold.

And prediction games may benefit players by motivating them to learn more about a domain or even develop a skill unrelated to the game. T-NA-5 posited that prediction games could be "*a good motivation for education.*" S/he described the inspiration for his/her game (G10, the game about ranking the contestants in a reading competition): "*So the setting that I design is...let's say my lab...we create a paper reading competition...to encourage students to read academic papers and enhance their knowledge. So during...the summer...students are required to log the papers that they read in a common spreadsheet...Because my lab was trying to do it last summer and my advisor created a spreadsheet, but no one put anything there.*"

Overall, the range of activities generated indicate that the potential variety of prediction games is broader than their original conceptualization based on fantasy sports. Furthermore, there are three factors that may be important in the design of an engaging prediction game: educational, social, and socio-cultural. Future work may investigate which combinations of these factors and their emphasis have an effect on engagement and learning.

14. CONCLUSION

14.1. Designing Prediction Games[†††]

Prediction games, data-driven games modeled after fantasy sports, are aimed to motivate people to explore, analyze, and develop their own understanding of large data sets. They revolve around activities where players examine historical data and information resources to make predictions about future events. As a result, they may help improve the players' domain knowledge and data interpretation skills. Experience with the Fantasy Forecaster prediction-game prototype found that players wanted to explore and analyze data and became more interested in and knowledgeable about the domain during game play. Results from a survey of fantasy sports users highlight that data gathering and analysis play a significant role for many fantasy sports players and that they frequently use data and data analysis tools external to the game environment. A frequently-reported category of external data was reading news and on-line commentary about the domain. The survey also identified the motivational effects of social ties to other players and the importance of communication between players.

Results from use of the Fantasy Forecaster prototype and the survey of fantasy sports players led to a framework describing the major prediction game components. These include the expected capabilities to archive and explore domain data, facilities for

[†††] Part of this section is reprinted with permission from the paper: © 2016 ACM. Gabriel Dzodom and Frank Shipman. 2016. Data-driven Prediction Games. In proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA' 16). Association for Computing Machinery, New York, NY, USA, 1857 – 1864. DOI:https://doi.org/10.1145/2851581.2892546

players to make and observe the outcome of predictions, and the collection of data from new observations with which to score predictions. But the experiences pointed to additional components important for player motivation and engagement. In particular, the framework emphasizes the role of readable/viewable domain-related information resources and methods for player communication and community building.

Our first full prediction game instantiation, Fantasy Climate, has players choose cities and predict which are the warmest and coolest relative to historic norms. While much of the gaming environment is domain independent, Fantasy Climate features a map-based prediction interface for players to pick their locations and a data visualization tailored to the decisions being made by players. Additionally, Fantasy Climate includes in-game news access and multiple methods for player communication, features found to be missing from many other fantasy-sports inspired games and thought to be important to maintaining player engagement.

User studies evaluated the news and communication components of our design as well as the experiences of prediction game players more generally. The first study examined the use and effects of four different communication tools: player-to-player chat and email as well as a temporally-structured comment wall and a topically-structured forum. Overall, players confirmed the importance of communication in keeping them engaged in the gaming activity. Of the available tools, the forum was favored by the majority of players and was the hub for many interactions among members during the activity. Forum topics often involved in-depth discussions of prediction techniques or data analysis methods. Another finding is that the asynchronous tools (the forum, the comment wall, and direct messages) had a more positive effect on engagement than did chat/IM in part due to the need for players to be in the game at the same time and the experimental setup where players did not know one another before the activity.

The two alternative interfaces for presenting domain-specific news, map-based GeoNews and traditional snippet-list Newsboard, were compared in the second study. Players with GeoNews reported greater engagement with the activity. This conclusion was supported by all metrics including the overall time spent on game activities, the average time spent per visit to Fantasy Climate, and the average time spent per news article accessed. Additionally, the quality of predictions made by players with GeoNews was better than it was for Newsboard and they improved over the course of the activity. Taken together, the findings indicated that the presentation of domain-related news has a strong effect on engagement and performance.

Overall, prediction games have the potential to encourage domain and data analysis skills and learning. Unlike fantasy sports, prediction games in other domains cannot rely on players' inherent interest in the domains. Gamification can provide motivation to players but other features of gaming environments play an important role in keeping players interested and active. Here we have identified player-to-player communication and domain-related news as important components of prediction games and evaluated their effects.

Several issues arose during the studies that deserve further exploration. In the first study, the lack of prior social connections between players was reported as inhibiting the use of some communication tools, particularly chat. How would the results vary for a group of players that have more knowledge of one another, such as the students in a class? This is an important use case envisioned in the design of prediction games that may result in different behaviors than those found in our results. In the second study, the player performance for the NewsBoard group actually declined over the course of the activity. Did they get bored and care less? Did the subsets of location and weather patterns make the task more difficult? A future study needs to gather player experiences and reactions over the course of the activity rather than only after the activity is complete. Finally, we have been exploring the presentation of additional domain-focused information, such as social media posts (e.g. tweets). Would such presentation be a valuable addition or just noise given the existing data and news content?

Our work has shown that prediction games have the potential to increase engagement with data and other content. Therefore, we envision prediction games in formal and informal educational settings. Inspired by the work on environments supporting design, this vision led to the exploration of an authoring environment that allows teacher or instructional designers to create prediction activities without writing a single line of code.

14.2. Authoring Prediction Activities

Using our prediction games framework and the lessons from the implementation of Fantasy Climate, we identified seven components and optional features that define a prediction activity. First are the name or title of the activity (e.g. “Extreme Locations”) and the directives to remind players of the activity objectives. Second is the prediction schedule that dictates when the activity starts, when it ends, and when the players' predictions are due and scored. Third are the selection sets which are sets around which players' predictions revolve. In Fantasy Climate, these are location (or U.S. city) sets. Fourth are the scoring rules, expressed as mathematical formulas, used to score the players' predictions. Fifth are the options for supporting prediction making like the temperature visualization tool Thermovizz developed for Fantasy Climate. Sixth are options to facilitate social interaction through competition or communication. And seventh are the activity members which are the users that will participate in the prediction activity. In a formal educational setting, students are expected to be the activity members and the instructor, the activity author.

We developed the Activity Creation Wizard (ACW), an environment that guides the user through a series of steps to author a prediction activity for Fantasy Climate. The ACW steps were informed by the seven components of a prediction activity. They are: (1) defining the activity objectives, (2) building the prediction schedule, (3) creating the selection sets, (4) defining the scoring rules, (5) selecting support for prediction making, (6) customizing community interactions, and (7) identifying activity members. The ACW also features a template component that allows the activity author to reuse the customizations of a previously created prediction activity, and tools for automatically generating the prediction schedule and the selection sets. To reduce the required background knowledge, the ACW incorporates a help system that provides explanations, examples and tutorials to the activity creator during the authoring process.

A user study was conducted to evaluate the ACW, to better understand the activity authoring process, and to find out the kind of prediction games that people may find engaging. Hence, the participants performed two tasks: (1) to author a prediction activity for Fantasy Climate given some requirements, and (2) to design a prediction game in the domain of their choosing.

14.2.1. ACW Evaluation

All 24 participants in the ACW's evaluation successfully authored a prediction activity in the climate domain. The help system was adequate in assisting the authors (e.g. via explanations, examples) during the task. Even though such information was very detailed according to many participants and the automatic display of it through popup dialogs was found to be obstructive to others.

Participants found the most difficult steps of the authoring process to be (a) building the prediction schedule (Step 2), (b) creating the selection sets (Step 3), and (c) creating the scoring rules (Step 4). Schedule building in the ACW was found difficult because it was tedious and repetitive, especially when prediction activities span weeks with several prediction rounds. The creation of selection sets and the scoring rules were difficult because their effects on the prediction activity were opaque to the participants. At least for creating selection sets, participants in the conditions with automation could overcome this issue by using the Selection Set Generator. However, there was no such support for generating scoring formula, nor were the examples in the explanations and the template presets sufficient to help the participants in Step 4. As a consequence, the creation of the scoring rules was the most difficult step of the ACW.

The open ended prediction game authoring task provided two lessons that challenged our perspective of the prediction activity authoring process. One was that authors designed activities where the steps are interdependent, not mostly discrete and independent. For example, a user may want to tie selection sets to the prediction schedule (e.g. coastal cities for the summer or southern cities for the fall), or tie the selection sets to scoring rules. In this scenario, authoring the prediction activity will require that Step 2 (building the prediction schedule), Step 3 (creating the selection sets), and Step 4 (creating the scoring rules) of the ACW be interdependent which our current design does not allow. The other lesson was that, in designing a prediction activity for informal settings, authoring may be a collaborative process where the stakeholders would negotiate on different customizations of the activity like the prediction schedule, the scoring rules, or what tools should be available to support prediction making.

While only used by a subset of the participants with access to them, the template component and the automated tools were valuable to those participants who did use them. Many participants who created the prediction schedule manually suggested that a tool like the Automatic Schedule Builder would have improved their experience. Support via templates and automation raise issues of control, where authors may choose not to use them due to fear of not being able to do what they want. Having better explanations of the editability of the generated aspects of the design may help but this is a common trade-off across software of all types.

These lessons and other issues raised by the participants have implications in the next iteration of the ACW. First, the ACW user interfaces need enhancements such rearranging components to improve cohesiveness, using appropriate signifiers and icons to reduce confusion, and making system requirements more transparent to the user during the authoring process. Second, the content (e.g. explanations, examples) of the help system needs to be revised to remove language that may be obscure to users with less technical background and to make the text more concise without losing valuable information. Also, the help system should make greater use of progressive disclosure, and include a search or similar functionality to enable quick access to needed information.

Third, the effects of the customizations on the prediction activity need to be more transparent to the activity author, especially for the selection sets step and the scoring rules step of the ACW. This needs to go beyond just providing more examples and presets with explanations of their effects. It needs to help authors understand why these examples and presets make a good prediction activity. One direction for future research is the design of a tool to help generate and evaluate scoring formulas. In one version of such a tool, the activity author could manually enter prediction values and the tool would display the resulting scores. This would help authors better understand the effects of the formulas and ensure that the formulas produce the intended results. Beyond this formula evaluator would be a component that simulates a prediction activity using current customizations on historical data and visualizes the distribution of potential scores. Such a tool would enable the activity author to see effects of their customizations on the prediction activity and use the results to iterate on their design.

Fourth, the ACW could be redesigned to handle scenarios where customizations in one step are coupled with customizations in another step. For example, in Step 2 of the ACW, the new interface could allow the author to create multiple sub-parts of a prediction schedule and name them (e.g. First Half of Summer, Second Half of Summer.) Then in Step 3, the activity author could link each created selection set to sub-parts of the prediction schedule. Similarly, in Step 4, the scoring formulas could be linked to selection sets or schedule sub-parts.

Finally, the ACW could include features to facilitate collaboration during the authoring process of a prediction activity. Collaborative design could occur in a range of situations as, even in a formal educational setting, the activity creator and the instructor may not be the same person, or there might be multiple instructors (e.g. teaching assistant and professor).

14.2.2. Creating a Prediction Game

We interviewed the participants on their created prediction games. The findings revealed that the participants (and their peers) were interested in prediction games that were grounded in their socio-cultural contexts. For example, one participant created a gold price prediction game because of the important and expansive role that gold plays in her/his culture. Another one created a Marvel movie box office prediction game because s/he and her/his peers are Marvel movie fans. Also, in the games created, whether in person or through social media (e.g. Facebook, Twitter), peers were a major source information that players might seek to make predictions. This mirrors fantasy sports. Fantasy sports are engaging because sports in general are an important part of the culture and, as found in our survey of fantasy sports players, peers were source of information. Also, the variety of activities created by the participants also broadened our original conceptualization of prediction games beyond fantasy sports. Historical data may not be required for prediction games. The data driving them need not be structured and may even be produced by the players themselves.

Finally, participants expressed a recognition of the educational benefits of prediction games that align with the primary motivations of our research. One benefit was about improving players' knowledge of the game's domain or a related domain. For example, players of the gold price prediction game may learn about international geopolitics since they affect gold prices. And the other benefit was about motivating players to learn more about a domain or even develop a skill only peripherally related to the game.

Future studies of prediction game design could focus more on the educational, social and socio-cultural factors in their design. Understanding which factors and combinations thereof affect expected engagement and learning could help with both the generation of designs available via templates and the explanations of designs and their components. Perhaps the integration of more information resources into the prediction games system, and the development and evaluation of prediction games into other domains will provide insights into these questions.

14.3. Overall Future Work

Ultimately, we see prediction games as part of a larger instructional strategy. Overarching questions that drive this research are can players improve their domain understanding and data analysis skills through prediction games? And what kind of scaffoldings or support teachers or instructional designers with no technical background need to create prediction activities using the ACW? A learning assessment that uses pre- and post-tests would help resolve the former question. For the latter question, we need to engage in an iterative design approach of the ACW that would involve instructors and instructional designers.

REFERENCES

1. Shaaron Ainsworth and Piers Fleming. 2006. Evaluating authoring tools for teachers as instructional designers. *Computers in Human Behavior* 22, 1: 131–148. https://doi.org/10.1016/j.chb.2005.01.010

2. Judd Antin and Elizabeth F. Churchill. 2011. Badges in social media: A social psychological perspective. *CHI 2011 Gamification Workshop Proceedings (Vancouver, BC, Canada, 2011)*. Retrieved September 11, 2015 from http://uxscientist.com/?sort=post_date&page=7

3. David Axe. 2013. Hacking the Drone War's Secret History. *WIRED*. Retrieved July 8, 2014 from http://www.wired.com/2013/05/drone-api/

4. Josh Blackman, Adam Aft, and Corey M. Carpenter. 2012. FantasySCOTUS: Crowdsourcing a Prediction Market for the Supreme Court. *NorthWestern Journal of Technology and Intellectual Property* 10: 125.

5. Sabrina Bresciani and Martin J Eppler. 2008. The Risks of Visualization. Retrieved July 23, 2014 from http://sabrinabresciani.com/publications/the_risks_of_visualization/

6. Andrew Borg Cardona, Aske Walther Hansen, Julian Togelius, and Marie Gustafsson Friberger. 2014. Open Trumps Data Game. *Proceedings of Foundations of Digital Games*.

7. C. Chen. 2005. Top 10 Unsolved Information Visualization Problems. *IEEE Computer Graphics and Applications* 25, 4: 12–16. https://doi.org/10.1109/MCG.2005.91

8. Nickolas W. Davis and Margaret Carlisle Duncan. 2006. Sports Knowledge is Power Reinforcing Masculine Privilege Through Fantasy Sport League Participation. *Journal of Sport & Social Issues* 30, 3: 244–264. https://doi.org/10.1177/0193723506290324

9. Joris Drayer, Stephen L. Shapiro, Brendan Dwyer, Alan L. Morse, and Joel White. 2010. The effects of fantasy football participation on NFL consumption: A qualitative analysis. *Sport Management Review* 13, 2: 129–141. https://doi.org/10.1016/j.smr.2009.02.001

10. Brendan Dwyer. 2011. The impact of fantasy football involvement on intentions to watch NFL Games on Television. *International Journal of Sport Communication* 4, 3: 375–396.

11. Brendan Dwyer and Yongjae Kim. 2011. For Love or Money: Developing and Validating a Motivational Scale for Fantasy Football Participation. *Journal of Sport Management* 25, 1: 70–83.

12. Brendan Dwyer, Stephen L. Shapiro, and Joris Drayer. 2011. Segmenting Motivation: An Analysis of Fantasy Baseball Motives and Mediated Sport Consumption. *Sport Marketing Quaterly* 20, 3: 129–137.

13. Gabriel S. Dzodom and Frank M. Shipman. 2014. Data-driven Web Entertainment: The Data Collection and Analysis Practices of Fantasy Sports Players. In *Proceedings of the 2014 ACM Conference on Web Science* (WebSci '14), 293–294. https://doi.org/10.1145/2615569.2615649

14. Jon Erdman. 2014. Fantasy Snowfall League 2014. *The Weather Channel*. Retrieved July 8, 2014 from http://www.weather.com/news/weather-winter/fantasy-snowfall-league-2014-20131230

15. Wei Fang. 2007. Using Google Analytics for Improving Library Website Content and Design: A Case Study. *Library Philosophy and Practice* 2, 9. Retrieved from http://unllib.unl.edu/LPP/fang.htm

16. Lee K. Farquhar and Robert Meeds. 2007. Types of Fantasy Sports Users and Their Motivations. *Journal of Computer-Mediated Communication* 12, 4: 1208–1228. https://doi.org/10.1111/j.1083-6101.2007.00370.x

17. Paula Felps. Fantasy sport players get a big assist from the Internet and new software. Retrieved September 21, 2016 from http://www.newson6.com/story/7677401/fantasy-sport-players-get-a-big-assist-from-the-internet-and-new-software

18. Thomas J. Fennewald and Brent Kievit-Kylar. 2013. Integrating Climate Change Mechanics Into a Common Pool Resource Game. *Simulation & Gaming* 44, 2–3: 427–451. https://doi.org/10.1177/1046878112467618

19. Stephen Few. 2007. Data Visualization Past, Present, and Future. Retrieved July 23, 2014 from http://www.perceptualedge.com/articles/Whitepapers/Data_Visualization.pdf

20. G. Fischer, A. Girgensohn, K. Nakakoji, and D. Redmiles. 1992. Supporting software designers with integrated domain-oriented design environments. *IEEE Transactions on Software Engineering* 18, 6: 511–522. https://doi.org/10.1109/32.142873

21. Gerhard Fischer. 1994. Domain-oriented design environments. *Automated Software Engineering* 1, 2: 177–203. https://doi.org/10.1007/BF00872289

22. Marie Friberger, Julian Togelius, Andrew Borg Cardona, Michele Ermacora, Anders Mousten, Martin Moller Jensen, Virgil-Alexandru Tanase, and Ulrik Brondsted. 2013. Data Games. *Proceedings of the Procedural Content Generation Workshop at Conference on the Foundation of Digital Games*.

23. M.G. Friberger and J. Togelius. 2012. Generating interesting Monopoly boards from open data. In *2012 IEEE Conference on Computational Intelligence and Games (CIG)*, 288–295. https://doi.org/10.1109/CIG.2012.6374168

24. Rohit Gargate. 2013. Weather Data Gamification. Texas A&M University, College Station.

25. Rosemary Garris, Robert Ahlers, and James E. Driskell. 2002. Games, Motivation, and Learning: A Research and Practice Model. *Simulation & Gaming* 33, 4: 441–467. https://doi.org/10.1177/1046878102238607

26. Matthew K. Gold. 2012. *Debates in the Digital Humanities*. Univ Of Minnesota Press, Minneapolis. Retrieved from http://www.amazon.com/Debates-Digital-Humanities-Matthew-Gold/dp/0816677956

27. Murray W. Goldberg, Sasan Salari, and Paul Swoboda. 1996. World Wide Web-course Tool: An Environment for Building WWW-based Courses. In *Proceedings of the Fifth International World Wide Web Conference on Computer Networks and ISDN Systems*, 1219–1231. Retrieved March 24, 2017 from http://dl.acm.org/citation.cfm?id=232710.232745

28. M. P. Jacob Habgood and Shaaron E. Ainsworth. 2011. Motivating Children to Learn Effectively: Exploring the Value of Intrinsic Integration in Educational Games. *Journal of the Learning Sciences* 20, 2: 169–206. https://doi.org/10.1080/10508406.2010.508029

29. Sandra Hirsh, Christine Anderson, and Matthew Caselli. 2012. The Reality of Fantasy: Uncovering Information-seeking Behaviors and Needs in Online Fantasy Sports. In *CHI ’12 Extended Abstracts on Human Factors in Computing Systems* (CHI EA ’12), 849–864. https://doi.org/10.1145/2212776.2212858

30. Young Ik Suh, Choonghoon Lim, Dae Hee Kwak, and Paul Pedersen. 2010. Examining the psychological factors associated with involvement in fantasy sports: An analysis of participants’ motivations and constraints. *International Journal of Sport Management, Recreation and Tourism* 5: 1–28. https://doi.org/10.5199/ijsmart-1791-874X-5a

31. Reddy Meghanath Junnutula. 2015. Asynchronous and synchronous communications’ effect on user engagement in prediction games. Texas A&M University, College Station.

32. Adam J. Karg and Heath McDonald. 2011. Fantasy sport participation as a complement to traditional sport consumption. *Sport Management Review* 14, 4: 327–346. https://doi.org/10.1016/j.smr.2010.11.004

33. Milos Kravcik, Marcus Specht, and Reinhard Oppermann. 2004. Evaluation of WINDS Authoring Environment. In *Adaptive Hypermedia and Adaptive Web-Based Systems*, 166–175. https://doi.org/10.1007/978-3-540-27780-4_20

34. Akshay Kulkarni. 2016. Effect of Visualization of News Articles in Data Driven Games. Texas A&M University, College Station.

35. Seunghwan Lee, Won Jae Seo, and B. Christine Green. 2013. Understanding why people play fantasy sport: development of the Fantasy Sport Motivation Inventory (FanSMI). *European Sport Management Quarterly* 13, 2: 166–199. https://doi.org/10.1080/16184742.2012.752855

36. Denise Lu. 2013. 7 Data Viz Sites to Inspire Your Creative Eye. *Mashable*. Retrieved July 22, 2014 from http://mashable.com/2013/10/01/data-viz-sites/

37. James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. 2011. *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute. Retrieved July 8, 2014 from http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation

38. Viktor Mayer-Schönberger and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Eamon Dolan/Houghton Mifflin Harcourt, Boston. Retrieved from http://www.amazon.com/Big-Data-Revolution-Transform-Think/dp/0544002695

39. Andrew McAfee and Erik Brynjolfsson. 2012. Big Data: The Management Revolution. *Harvard Business Review*. Retrieved July 8, 2014 from http://hbr.org/2012/10/big-data-the-management-revolution/ar

40. R. McDaniel, R. Lindgren, and J. Friskics. 2012. Using badges for shaping interactions in online learning environments. In *Professional Communication Conference (IPCC), 2012 IEEE International*, 1–4. https://doi.org/10.1109/IPCC.2012.6408619

41. Thomas L. Naps, Guido Rößling, Vicki Almstrum, Wanda Dann, Rudolf Fleischer, Chris Hundhausen, Ari Korhonen, Lauri Malmi, Myles McNally, Susan Rodger, and J. Ángel Velázquez-Iturbide. 2002. Exploring the Role of Visualization and Engagement in Computer Science Education. In *Working Group Reports from*

ITiCSE on Innovation and Technology in Computer Science Education (ITiCSE-WGR ’02), 131–152. https://doi.org/10.1145/782941.782998

42. Todd M. Nesbit and Kerry A. King. 2010. The Impact of Fantasy Football Participation on NFL Attendance. *Atlantic Economic Journal* 38, 1: 95–108. https://doi.org/10.1007/s11293-009-9202-x

43. Stephanie Norman. 2016. 10 Authoring Tools For Easy eLearning Design - eLearning Industry. *eLearning Industry*. Retrieved March 24, 2017 from https://elearningindustry.com/10-authoring-tools-easy-elearning-design

44. Beatriz Plaza. 2011. Google Analytics for measuring website performance. *Tourism Management* 32, 3: 477–481. https://doi.org/10.1016/j.tourman.2010.03.015

45. Marc Prensky. 2005. Computer games and learning: Digital game-based learning. In *Handbook of Computer Game Studies*. MIT Press, 97–122. Retrieved July 24, 2014 from http://www.itu.dk/people/jrbe/DMOK/Artikler/Computer%20games%20and%20learning%202006.pdf

46. Quint Randle and Rob Nyland. 2008. Participation in Internet Fantasy Sports Leagues and Mass Media Use. *Journal of Website Promotion* 3, 3–4: 143–152. https://doi.org/10.1080/15533610802077180

47. B. J. Ruihley and R. Hardin. 2013. Meeting the informational needs of the fantasy sport user. *Journal of Sports Media* 8, 2: 53–80. https://doi.org/10.1353/jsm.2013.0013

48. Brody J. Ruihley and Andrew C. Billings. 2013. Infiltrating the boys’ club: Motivations for women’s fantasy sport participation. *International Review for the Sociology of Sport* 48, 4: 435–452. https://doi.org/10.1177/1012690212443440

49. Brody J. Ruihley and Robin L. Hardin. 2011. Beyond touchdowns, homeruns, and three-pointers: an examination of fantasy sport participation motivation. *International Journal of Sport Management and Marketing* 10, 3/4: 232–256. https://doi.org/10.1504/IJSMM.2011.044792

50. Brody J. Ruihley and Robin L. Hardin. 2011. Message Boards and the Fantasy Sport Experience. *International Journal of Sport Communication* 4, 2: 233–252.

51. Jeff Sauro. 2011. MeasuringU: Measuring Usability with the System Usability Scale (SUS). *MeasuringU: Measuring Usability with the System Usability Scale (SUS)*. Retrieved August 1, 2020 from https://measuringu.com/sus/

52. Jeff Sauro. 2011. SUStisfied? Little-Known System Usability Scale Facts User Experience Magazine. *SUStisfied? Little-Known System Usability Scale Facts User Experience Magazine*. Retrieved August 1, 2020 from http://uxpamagazine.org/sustified/

53. Frank M. Shipman. 2009. Blending the Real and Virtual in Games: The Model of Fantasy Sports. In *Proceedings of the 4th International Conference on Foundations of Digital Games* (FDG ’09), 169–174. https://doi.org/10.1145/1536513.1536547

54. Brian Smith, Priya Sharma, and Paula Hooper. 2006. Decision making in online fantasy sports communities. *Interactive Technology and Smart Education* 3, 4: 347–360. https://doi.org/10.1108/17415650680000072

55. Will T. 2017. Measuring and Interpreting System Usability Scale (SUS). *UIUX Trend*. Retrieved August 1, 2020 from https://uiuxtrend.com/measuring-system-usability-scale-sus/

56. Kerry King Todd Nesbit. 2010. The Impact of Fantasy Sports on Television Viewership. *Journal of Media Economics* 23, 1: 24–41. https://doi.org/10.1080/08997761003590721

57. Julian Togelius and Marie Gustafsson Friberger. 2013. Bar Chart Ball, a Data Game. *Proceedings of Foundations of Digital Games*: 451–452.

58. Katrien Verbert, Erik Duval, Joris Klerkx, Sten Govaerts, and José Luis Santos. 2013. Learning Analytics Dashboard Applications. *American Behavioral Scientist*: 0002764213479363. https://doi.org/10.1177/0002764213479363

59. Donghee Yvette Wohn, Emma J. Freeman, and Katherine J. Quehl. 2017. A Game of Research: Information Management and Decision-making in Daily Fantasy Sports. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (CHI PLAY ’17), 355–366. https://doi.org/10.1145/3116595.3116605

60. 2013. System Usability Scale (SUS). Retrieved August 1, 2020 from system-usability-scale.html

61. Data.gov. *Data.gov*. Retrieved July 8, 2014 from https://www.data.gov/

62. Dronestream. *Dronestream*. Retrieved July 8, 2014 from http://dronestre.am

63. Out of Sight, Out of Mind. *Out of Sight, Out of Mind*. Retrieved July 8, 2014 from http://drones.pitchinteractive.com

64. United States Gun Killings. Retrieved July 22, 2014 from http://guns.periscopic.com

65. Beautiful News. *Beautiful News*. Retrieved July 14, 2020 from https://informationisbeautiful.net/beautifulnews/q/topic:eco-and-climate/

66. Fantasy is Reality: A Look at the Growing Engagement in Fantasy Sports. Retrieved July 14, 2020 from https://www.nielsen.com/us/en/insights/article/2018/fantasy-is-reality-a-look-at-growing-engagement-in-fantasy-sports

67. About FantasySCOTUS. *FantasyScotus From the Harlan Institute*. Retrieved July 9, 2014 from http://www.fantasyscotus.net/about-fantasyscotus/

68. Fantasy Film League. *Fantasy Film League*. Retrieved July 8, 2014 from http://www.fantasyfilmleague.com/

69. Summer Movie League. *summermovieleague.com*. Retrieved July 8, 2014 from http://summermovieleague.com/Welcome.aspx

70. Welcome to Fantasy Congress. *Fantasy Congress*. Retrieved July 8, 2014 from http://www.fantasycongress.net/112/index.php

71. Predict It. *Predict It @ predict.realius.com*. Retrieved July 8, 2014 from http://predict.realius.com/

72. Wall Street Survivor. *Wall Street Survivor*. Retrieved July 8, 2014 from http://www.wallstreetsurvivor.com

APPENDIX A

QUESTIONNAIRE ON THE PRACTICES AND EXPERIENCE OF FANTASY SPORTS USERS

The questionnaire combines multiple choice questions, 7-point Likert-scale questions, and open-ended questions. Sets of related Likert-scale questions with follow-up open ended questions were used to gage the participants' sentiments or practices regarding each aspect of the game.

In the Likert-scale questions below, sentiment options = (Strongly Disagree, Disagree, Somewhat Disagree, Neutral, Somewhat Agree, Agree, Strongly Agree).

For skill level questions, options = (None, Basic, Novice, Intermediate, Advanced, Master, Expert).

For validation questions, options = (Never True, Almost Never True, Usually Not True, Occasionally True, Usually True, Almost Always True, Always True).

1. *What is your age?*
    - 18 – 25
    - 26 – 40
    - 41 – Older

2. *What is your gender?*
    - Male
    - Female
    - Other

3. *How would you classify yourself?*
    - Asian/Pacific Islander

- Black
- Caucasian/White
- Hispanic/Latino
- Multiracial
- Other
- Would rather not say

4. *What is the highest level of education you have completed?*
- High school or equivalent
- Vocational or equivalent
- Some college
- Bachelor degree
- Master degree
- Doctoral degree
- Professional degree (MD, JD, etc...)

5. *What best describes your field of study in college?*
- Liberal Arts
- Social Sciences
- Natural Sciences
- Engineering

6. *During your studies did you take any course in*	Yes	No
statistics		
probability		
stochastic processes		
modeling and optimization		

7. *Rate your knowledge of*	Skill level options						
statistics							
probability							
stochastic processes							
modeling and optimization							

8. *How long have you been playing fantasy sports game?*
- Less than 6 months
- About a year
- 1 – 2 years

- More than 2 years

9. *During last season (or the current one), how long did the average interaction with the game last?*
- Less than 15 mins
- 15 mins – 30 mins
- 30 mins – 1 hour
- More than 1 hour

10. *How many interactions (with the game) per week did you have during last season?*
- 1
- 2 - 4
- 5 - 7
- More than 7

11. *You play fantasy sports because*	Sentiment options						
it is another fun way to engage with and understand sports							
it is just good entertainment like any other game							
you can make money while having fun							
it is a great way to spend time with your friends who also play fantasy sports							
you like the intellectual challenge and skill it takes to select the right team to win							

12. *Please list other reasons(not mentioned above) why you play fantasy sports online*

13. *Besides fantasy sports, please list describe other sports related activities have you recently engaged with (i.e. attending games, video games, visiting sports websites, etc...)*

14. *During last(or current) season gameplay, you spent more time*	Sentiment options						
creating/joining a league							
reading, and analyzing teams/players statistics							

selecting players and building teams							
managing your league and teams(ie: trading players)							
checking your score, the game status							

15. *Please list and describe any other time consuming activity not mentioned in the previous question*

16. *During last (or current) season, how many active leagues were you a part of?*
- 0
- 1 – 2
- 3 – 4
- More than 4

17. *Please elaborate on your reasons for joining these leagues*

18. *Please describe your process for selecting players*

19. *During your last team/players selection, you used*	Sentiment options						
in-game resources (news feed, forums, message boards, chats)							
external online resources (Social Networks, Sport News website, external forums & message boards)							
offline resources(Magazine, Television, Friends)							

20. *Please list and describe other resources (not mentioned in the previous question) used for player/team selection*

21. *During your last analysis of team/player statistics and information*	Sentiment options						
You only used the tools supported by the fantasy game							
You used a combination of game-supported tools and external tools (i.e. excel, a sheet of paper, etc...)							
You only used the external tools (i.e.: excel, a sheet of paper, etc...)							

22. *Please elaborate on the reasons why you preferred game-supported tools, external tools, or a combination for team/players data analysis*

23. *During your last analysis of team/player statistics and information you used*	Sentiment options						
your formal background in either of the fields(Statistics, Probability, Stochastics, or Modeling and Optimization)							
just basic mathematics							
guesses and deductions							

24. *During your last analysis of team/player statistics and information*	Sentiment options						
you found the volume of data in the fantasy sport game overwhelming							
the data presentation and organization in the fantasy sport game made it difficult to extract useful information							
you would have liked more control over data presentation, and organization in the game							
data layout and presentation in the game was not important to you							

25. *Please list any other game constraints (not mentioned in the previous question) or lack of features that has hindered your data analysis activities*

26. *If your decisions depended on data analysis, how many variables (i.e.: speed, scoring points, passing points, etc...) did you consider in order to select a player*

- 0
- 1 - 4
- 5 - 7
- 8 - 10
- More than 10

27. *After you have made your selections for a game, how often have you logged back in the game to update your selections (changing players based on new information, trade players, etc...)?*

- Never

- Rarely
- Occasionally
- Frequently
- Very Frequently

28. *After you have made your selections*	Validation options						
you use the in-game tools to simulate a game with your selections to ensure they are optimal							
you would like to simulate a game with your selections to ensure they are optimal, but the fantasy game does not support such feature, and if it does it is not gratis							
you forget about the game until you get a notification about the results							
you frequently visit the fantasy game or other resources to get the latest updates on your teams/players							

29. Please describe any other interactions (not mentioned in the previous question) that you have had with the game after you have made your selections

30. Please describe all interactions you have had with the fantasy sport game in between seasons (i.e.: no interaction, social (forums, Chats, etc...), etc...)

31. Please describe any feature you think will improve your experience with Fantasy Sports.

APPENDIX B

PREDICTION GAMES SYSTEM ARCHITECTURE

Like fantasy sports, the prediction game environment is built as a combination of web applications. The system architecture employs the standard service-oriented approach with three primary layers: (a) the client which is the set user interfaces, (b) the web service which maintains the logic of the prediction game engine, and (c) the storage to manage the game content. These layers are independent of each other and communicate through application programming interfaces (APIs). This way, changes or extensions to the system are well contained and virtually non-disruptive to the cohesion of the environment. Figure 73 depicts the system architecture.

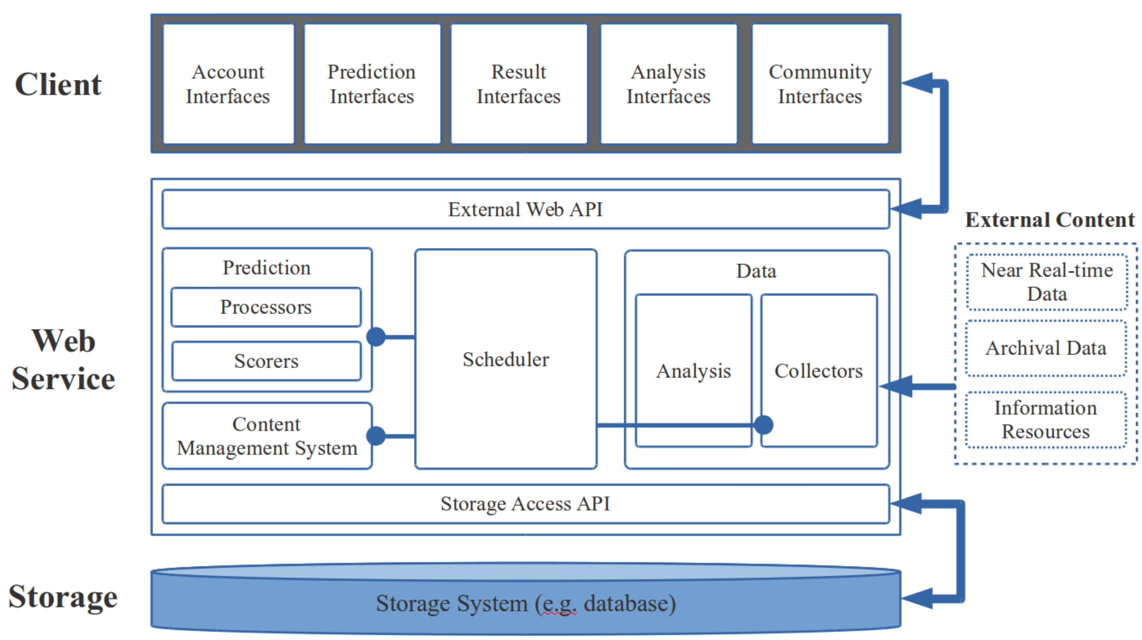

Figure 73. Prediction games system architecture

The client layer is the web portal through which users interact with the prediction game system. Through the account interfaces, players create or update a user account and authenticate (i.e. login) to the prediction game website. During an ongoing activity, the players may access the prediction interface to make and submit their prediction. The prediction interface varies depends on the particular prediction game being played. For example, the prediction games Fantasy Forecaster and Fantasy Climate, although in the same domain and based on the same data (weather), have different prediction interfaces. The analysis interfaces are a set interactive data and information presentation (e.g. visualization) interfaces to aid prediction making. Once predictions have been scored, players may view their performance and rank via the result interfaces. And, the community interfaces facilitate social interaction during the activity. A web API bridges the client and the web service. The client sends web requests to the web service and receives a response in a standard notation (e.g. JSON). Thus, the client is extensible and permit third-party add-ons as long as one adheres to the specifications of the web API specifications.

The web service envelops the prediction game engine. It mainly serves the requests from the client. For example, a data presentation interface would request a subset of historical domain data for its visualization component, while a message board user interface would poll for the latest players' messages. However, the web service also accesses the storage via an API. This storage-agnostic API is a collection of methods that specify the common operations (Create, Retrieve, Update, and Delete) allowable on the storage. This way, replacing the storage system only requires the API implementation of the new storage mechanism.

Since prediction activities have a duration and a prediction schedule, the web service features one extra component, the scheduler, in the predication game engine that handle components that depend on the activity schedule. The scheduler starts and ends the prediction activity at the start date and the end date respectively, as configured by the activity author. During the activity, it sends notifications to players to remind them of the predictions' submission deadlines. It also periodically invokes (as configured by the developers) the information resources collector to get the latest domain-related information (e.g. news, tweets, etc...). On a scoring date, the scheduler invokes the real-time data collector to retrieve the domain data for scoring. Then, it invokes the prediction processor to transform the players' predictions into numerical values. The execution of the prediction scorer follows right away with the transformed predictions and the retrieved domain data. Once the predictions' scores are computed, the scheduler notifies the players, inviting them back to the game site to view their performance. Thus, the scheduler web service is the brain of the prediction game system.

Finally, the storage is simply the structured container of the game content coming either from the users, or from the domain data and information resources. In the current implementation, the storage is a SQL database management system.

APPENDIX C

FANTASY PRECIPITATION

Like Fantasy Climate, Fantasy Precipitation is also based on weather data but focuses on precipitation. The goal is to engage people with precipitation patterns around a region (e.g. U.S.) and how such patterns have varied over time. Fantasy Precipitation encourages the players to observe long term frequency in precipitation data, and employ such observations to do well during the game. Hence players are asked to predict whether precipitation (snow or rain) will occur or not within a date range (e.g. 7 days) for every city of a given set. Scoring is based on the number of correct predictions. For example, a score of **4 / 5** in a prediction round means that the player made correct predictions for four cities and an incorrect prediction for a single city. Finally, Fantasy Precipitation support players in their prediction making activities by providing visualizations for precipitation data (including historical), and for streamed news articles related to weather.

Fantasy Precipitation and Fantasy Climate share the same domain and dataset (weather data). As a consequence, Fantasy Precipitation is also built using the prediction game system described in Appendix B and shares most of its system components with Fantasy Climate. The implementation of Fantasy Precipitation has primarily involved the development of a new prediction interface along with a new method to process predictions submitted by the players, a new data analysis tool to assist the players in their decision making, and a new results interface for players' scores. Also, there have been minor user interface changes to reflect the new game (e.g. updating all the 'Fantasy Climate' labels to 'Fantasy Precipitation'). However, these last minor modifications are not significant enough to warrant further elaboration.

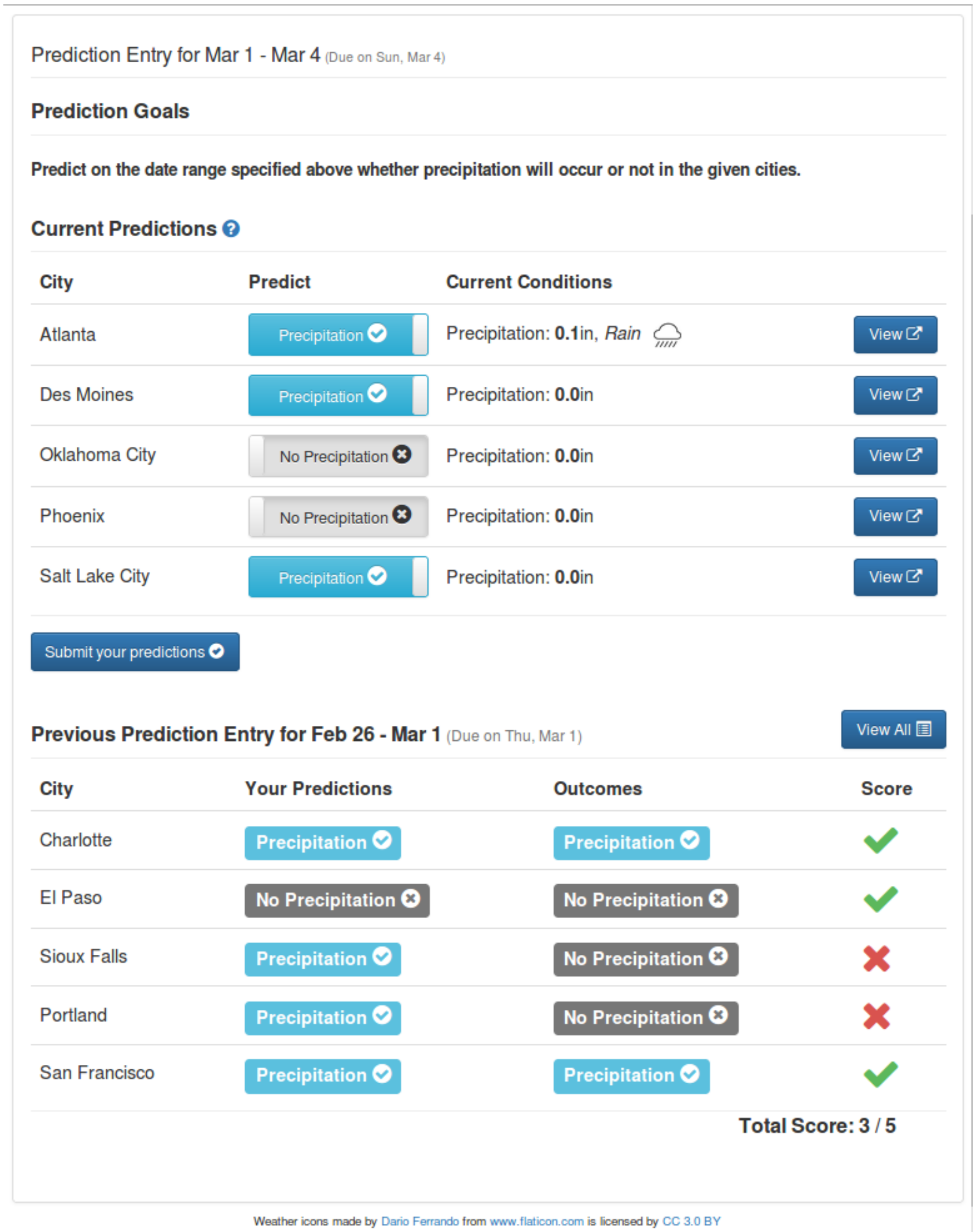

Figure 74. Fantasy Precipitation prediction interface

The new prediction interface in Fantasy Precipitation is structurally similar to Fantasy Climate's. At the top is a description of the players' directives or prediction goals. The next component enables the players to input their predictions. For every location of a prediction round, players predict whether precipitation will occur or not by toggling the switch control next to the location's name. For example, in Figure 74, the player is predicting that, in the period between March 1st and March 4th, there will be precipitation in Atlanta, Des Moines, and Salt Lake City but no precipitation in Oklahoma City and Phoenix. Adjacent to the switch control are the current precipitation conditions along with a button to launch the Location Profile tool shown in Figure 75. This tool provides a current radar image, related news, and data about prior years' precipitation totals for the period (dates) of the prediction.

The bottom of the prediction interface shows the results of the player's previous predictions. For every location, the players' prediction is displayed accompanied by its corresponding outcome and score. A green check signifies a correct prediction (prediction matches the outcome) and, a red X signifies otherwise. In Figure 74, the player made three correct predictions out of five hence his score: **3 / 5**. Finally like in Fantasy Climate, players can navigate to an interface that lists all their past predictions with their scores by pressing the 'View All' button.

The Location Profile tool assist the players in their prediction making. The interface shown in Figure 75 below comprises three components for every location in the prediction round: a weather map, a short list of weather-related news focusing on precipitation, a bar chart of historical precipitation data. The weather map is layered with precipitation pattern visuals. It has controls (above the map) that enables the players to observe precipitation patterns at a future date up to 12 days ahead. The bar chart visualizes the yearly precipitation pattern from 2010 to 1970. It graphs the average (e.g. 4 days average) of precipitation accumulation (in inches) over the years for the designated period of the current prediction round. Figure 75 shows the yearly precipitation average for period Mar 1 – Mar 4 about Atlanta, Georgia.

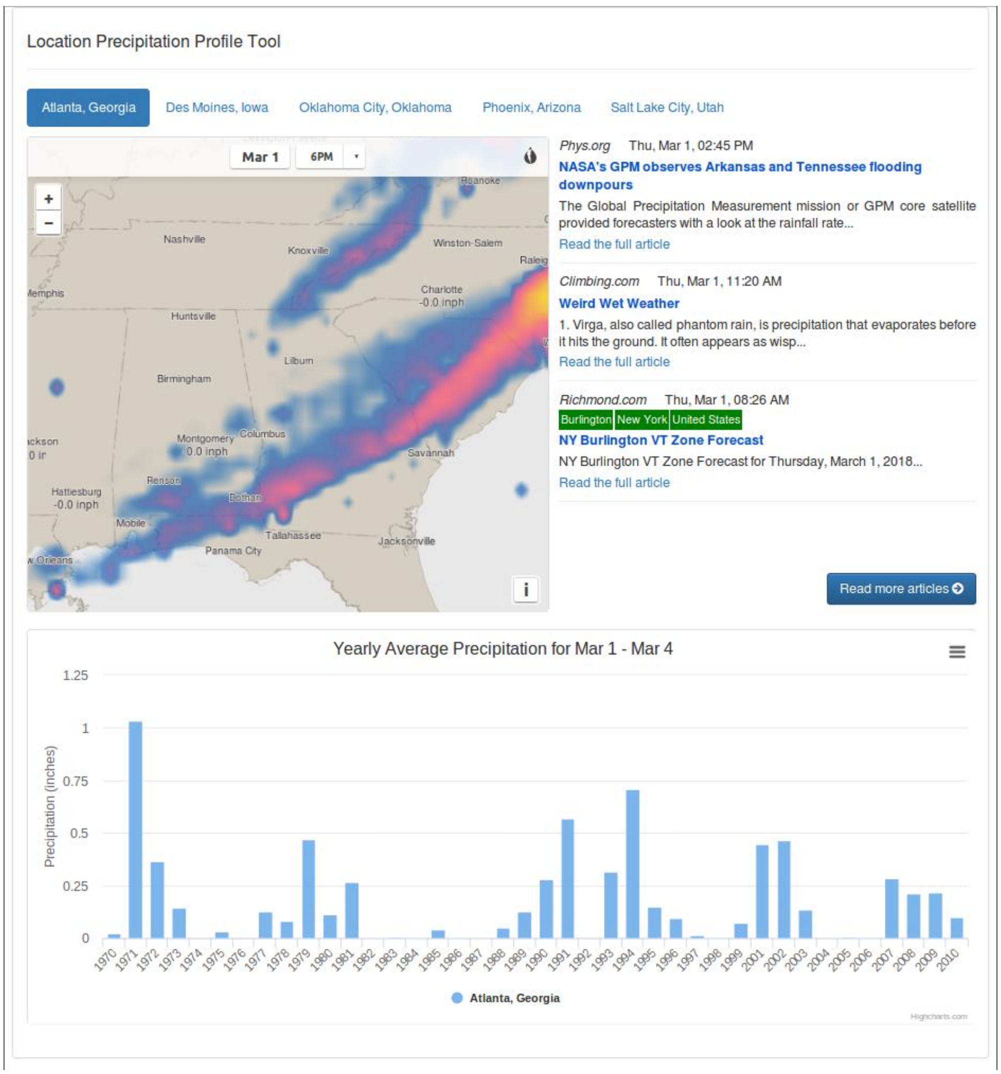

Figure 75. Location Profile tool for Fantasy Precipitation

APPENDIX D

USER STUDY CONSENT FORM

You are invited to take part in a research study being conducted by Gabriel Dzodom, a graduate student at Texas A&M University under the direction of Dr. Frank Shipman. The information provided in this form is to help you decide whether or not to take part in the study. If you decide to do so, you will be asked to sign this consent form. On the other hand, if you do not want to participate, there will be no penalty to you, no loss of any benefits.

Why Is This Study Being Done?

Prediction games are data-driven games modeled after fantasy sports where players examine data and other information resources to make predictions about future events. The study seeks to identify the requirements of a system that enables the generation and customization of prediction games in other domains such as weather.

Why Am I Being Asked To Be In This Study?

There are no specific selection criteria to be able to participate in this study. You have the choice whether or not to be in this research study. However, you must be at least 18 years old or older to participate.

How Many People Will Be Asked To Be In This Study?

About 30 people will be invited to participate in this study.

What Are the Alternatives to being in this study?

The alternative to being in the study is not to participate.

What Will I Be Asked To Do In This Study?

Your participation will take about 120 minutes:

1. you will fill out a short questionnaire that tells us a little about you (e.g. your age, experience with fantasy sports, etc...);
2. you will watch a live demo of a game called Fantasy Climate that lets users predict the weather. You will be able to play the game yourself, and ask questions about it;
3. Fantasy Climate can be customized. You will be asked to customize it by following some straightforward instructions. Then we will ask you to dream up a prediction game that you would like to play;
4. you will answer another short questionnaire to give us feedback about the app you just used. Then you will have a short conversation with the research team about your experience.

Are There Any Risks To Me?

There are no foreseeable risks.

Will There Be Any Costs To Me?

Aside from your time, there are no costs for taking part in the study.

Will I Be Paid To Be In This Study?

You will receive a $20 gift card for your participation.

Will Information From This Study Be Kept Private?

No personally identifiable information will be collected for the study. The records will be kept private. Research records will be stored securely and only investigators (Gabriel

Dzodom, Frank Shipman) will have access to the records. Information about you and related to this study will be kept confidential to the extent permitted or required by law. People who have access to your information include the Principal Investigator and research study personnel. Representatives of regulatory agencies such as the Office of Human Research Protections (OHRP) and entities such as the Texas A&M University Human Subjects Protection Program may access your records to make sure the study is being run correctly and that information is collected properly.

Who may I Contact for More Information?

For more information, contact Gabriel Dzodom at gabriel.dzodom@email.tamu.edu or, the Principal Investigator Frank Shipman at shipman@cse.tamu.edu. For questions about your rights as a research participant, to provide input regarding research, or if you have questions, complaints, or concerns about the research, you may call the Texas A&M University Human Research Protection Program office by phone at 1-979-458-4067, toll free at 1-855-795-8636, or by email at irb@tamu.edu.

What if I Change My Mind About Participating?

Your participation is voluntary. You may decide to not begin or to stop participating at any time. If you choose not to take part in this study or, withdraw from it, there will be no consequence on your student status, medical care, employment, evaluation, relationship with Texas A&M University, etc.

STATEMENT OF CONSENT

I agree to be in this study and know that I am not giving up any legal rights by signing this form. The procedures, risks, and benefits have been explained to me, and my questions have been answered. I can ask more questions if I want. A copy of this entire consent form will be given to me.

APPENDIX E

USER STUDY RECRUITMENT EMAIL

Call for Participants,

Ever wanted to create your own online game? Please come and help us evaluate authoring tools for prediction games. They are games modeled after online fantasy sports where players use news and statistics to make predictions about future events. The study seeks to identify the requirements of a system that enables the generation and customization of prediction games in other domains such as weather.

Eligibility

You must be at least 18 years of age to participate.

When and Where

Your participation will take about 120 minutes. We will schedule an appointment based on your availability.

Location: Room 408G, Harvey R Bright Building (map)

What you will be doing

1. You will fill out a short questionnaire that tells us a little about you (e.g. your age, experience with fantasy sports, etc...);
2. you will watch a live demo of a game called ?Fantasy Climate? that lets users predict the weather. You will be able to play the game yourself, and ask questions about it;

3. Fantasy Climate can be customized. You will be asked to customize it by following some straightforward instructions. Then we will ask you to dream up a prediction game that you would like to play;
4. you will answer another short questionnaire to give us feedback about the app you just used. Then you will have a short conversation with the research team about your experience.

You will receive a $20 gift card for your participation.

How to sign up

If you are interested in participating, please email Gabriel at gabriel.dzodom@email.tamu.edu to schedule an appointment. For more information, email Gabriel Dzodom at gabriel.dzodom@email.tamu.edu or Professor Frank Shipman at shipman@cse.tamu.edu. For questions about your rights as a research participant; or if you have questions, complaints, or concerns about the research, you may call the Texas A&M University Human Subjects Protection Program office by phone at (979) 458-4067, toll free at 1-855-795-8636 or by email at irb@tamu.edu.

IRB Number: IRB2018-1491M   IRB Approval: 05/23/2019

Thank you for your time.

APPENDIX F

USER STUDY TASK 1 DESCRIPTION

Create and customize a prediction game activity for Fantasy Climate where players are rewarded for their good predictions. Players will make two selections per prediction round from a set of locations: (1) the location whose daily high temperature deviates the most from its historical norm and, (2) the location whose daily low temperature deviates the most from its historical norm. Thus, a good prediction would be of a location whose temperature (high/low) deviates greatly from its historical norm compared to other locations. Here are the requirements:

- name the activity Fantasy Climate Ex;
- the activity must be at least a semester long (~14 weeks);
- the predictions must be due biweekly (every 2 weeks) on Friday evenings and scored on Sunday evenings until the end of the semester;
- create a scoring system so that good predictions yield higher scores than poor ones. This means creating two formulas: one for high temperatures and one for low temperatures;
- players should have access to the data analysis tool Thermovizz but with the regression line visualization turned off;
- players should have access to news about the long term effect of temperature changes on the climate;
- players should be able to communicate with one another as they compete.

APPENDIX G

USER STUDY TASK 2 DESCRIPTION

Now that you’ve customized our game, try thinking up your own. You can work on the computer (in a text editor or sketch tool) or on paper to describe this new game. Imagine an aspect of a domain (other than climate or sports) you and your friends care about where prediction games could be applied: e.g. politics – who will win the 2020 presidential election, entertainment – who will be in the finale of Top Chef, etc...

- What will players predict? Remember that it needs to be something that can be measured like rank, rating, price, position, etc...
- How will players play the game? Use the type of measure you just picked to describe how players will compete. How will the winner be determined?
- Briefly describe your scoring system (including its justification)
- How will the players get or find the information or data they may need to make good predictions?

APPENDIX H

USER STUDY PRE-QUESTIONNAIRE

1. *What is your age?*

____

2. *What is your gender?*

____

3. How would you classify yourself?

- Asian / Pacific Islander
- Black / African American
- Caucasian / White
- Hispanic / Latinx
- Multiracial
- Other
- Would rather not say

4. What is the highest level of education you have completed?

- High school education
- Vocational training
- Some college
- Bachelor’s degree
- Master’s degree
- Doctoral degree
- Professional degree (MD, JD, etc...)

5. Which field most closely describes your major, training, or profession?

- Humanities (e.g. History, Literature, Philosophy, Art, etc...)
- Social Sciences (e.g. sociology, psychology, Economics, etc...)
- Physical Sciences (e.g. Physics, Mathematics, Chemistry, etc...)
- Engineering
- Architecture or Design
- Business
- Other: ____

6. What best describes your knowledge of statistics or data analysis?

- I have no knowledge of statistics or data analysis
- I have a basic knowledge of statistics or data analysis
- I have an average knowledge of statistics or data analysis
- I have an advanced knowledge of statistics or data analysis
- I have an expert knowledge of statistics or data analysis

7. What best describes your familiarity with fantasy sports?

- I have never played before
- I have played rarely or once in a while
- I have played occasionally
- I have played frequently
- I play all the time (every season)

8. What fantasy sport do you play the most?

- None
- Football
- Basketball
- Baseball
- Soccer

- Other: _______________

9. What best describes your familiarity with other online/web games?

- I have never played before
- I have played rarely or once in a while
- I have played occasionally
- I have played frequently
- I play all the time

APPENDIX I

USER STUDY POST-QUESTIONNAIRE

For the Likert-scale questions, sentiment options = (Strongly Disagree(SD), Disagree(D), Neutral(N), Agree(A), Strongly Agree(SA)).

	SD	**D**	**N**	**A**	**SA**
1. I think that I would like to use the ACW frequently					
2. I found the ACW unnecessarily complex					
3. I thought the ACW was easy to use					
4. I think I would need the support of a technical person to be able to use the ACW					
5. I found the various functions in the ACW were well integrated					
6. I thought there was too much inconsistency in the ACW					
7. I would imagine that most people would learn to use the ACW very quickly					
8. I found the ACW very cumbersome to use					
9. I felt very confident using the ACW					
10. I needed to learn a lot of things before I could get going with the ACW					

	SD	**D**	**N**	**A**	**SA**
11. I could have created and customized the game activity without any template					

12. The templates were helpful for creating a game activity.					

	SD	D	N	A	SA
13. The help system (?) icons/buttons were easy to find.					
14. The explanations / examples from the help system were easy to read and understand					
15. I found the automatic popup dialogs of explanations / examples from the help system annoying					
16. The explanations / examples from the help system were not helpful					

	SD	D	N	A	SA
17. Creating the prediction schedule for the game activity was tedious					
18. I found the automatic schedule builder easy to use					

	SD	D	N	A	SA
19. Creating the selection sets for the game activity was easy					
20. I found the selection sets generator difficult to use					

APPENDIX J

USER STUDY TASK 1 INTERVIEW QUESTIONS

Please describe your overall experience with the ACW

Do you think one needs an online games background to use the ACW easily to complete the task? What about a Fantasy Sports background?

What step or steps of the ACW were the most difficult for you? And Why? What do you think about the order of the ACW steps during the customization of the game activity? (e.g. were there steps you would have preferred to tackle first?)

What do you think about the explanations and examples dialogs? What did you have trouble with? What would have you helped you?

How would you describe your experience configuring the scoring formula for the game activity?

How would you describe your experience configuring the prediction schedule for the game activity?

How would you describe your experience configuring the selection sets for the game activity?

What do you think are the overall limitations of ACW? Are there aspects of Fantasy Climate or Prediction Games in general you would have love to customized but were not supported by the ACW? Are there aspects of Fantasy Climate or Prediction Games in general you would have love to customized that were supported by the ACW but were limited, incomplete or difficult to use?

APPENDIX K

USER STUDY TASK 2 INTERVIEW QUESTIONS

Can you briefly describe the motivations behind your design?

Who do you envision would be the audience for your game?

Do you think the current version of the ACW can help you create and customize your game? Why?

Anything you'd hoped I would ask you about that I didn't?