Dynamic Visualizations for Soccer Statistical Analysis

Adrian Rusu[1], Doru Stoica[2], Edward Burns[1], Benjamin Hample[1], Kevin McGarry[1], Robert Russell[1]
[1]Department of Computer Science, Rowan University
[2]Department of Physical Education, University of Craiova, Romania
rusu@rowan.edu, doru_stoica_2005@yahoo.com,
{burnse15, hample04, mcgarr92, russel42}@students.rowan.edu

Abstract

The analysis of large sums of soccer statistics can be extremely difficult if the data is not presented graphically. As a result, we developed an application, titled Soccer Scoop, which provides two separate visualizations that can aid a soccer team manager. With our application, a team manager can compare two players on different teams, analyze a particular player before signing them to a contract, measure the performance of a particular player at different positions, generate practice exercises, and determine if a particular player plays better on the road or at home. The visualizations used in our application apply information visualization techniques, such as glyphs, modified star plots, details on demand, color, and gestalt principles.

1. Introduction

In athletics, it is difficult for a team manager to determine the effectiveness of his or her players, especially when considering team specific pregame objectives such as the creation of drills and collaborative, tactical set plays. A team manager can maintain spreadsheets of useful data values, such as goals, team's average running time per game, and/or a player's speed at a given distance, however establishing a way to analyze each aspect of data as a distinct entity related to the overall team structure is one that may pose challenges. This information is vital to the success of a team, but unfortunately most team managers fail to utilize this information at its full extent.

Our application, called Soccer Scoop, attempts to solve this dilemma and to fulfill the need for player-to-player comparisons as related to team dynamics. Soccer Scoop presents two separate and very unique visualizations for soccer; all visualized by a quick observation technique that is made possible by establishing a soccer field metaphor [1] and using all the nature objects of the field to represent various data categories cross referenced with specific, key player elements. The goal is to create a visualization that minimizes the amount of cognitive effort required to understand it. By using metaphors, the visualization becomes appealing to the team manager [13], [14]. The first visualization, called Field Viewer, shows the user vital information about one player. This visualization represents the player's attributes in the form of a soccer field and with various entities on it. The second visualization is called the Player Viewer. The user is able to compare two players with this visualization, as each player is color coded to show the player's effectiveness in a given category.

From a team manager's perspective, in order to improve their roster, he or she wants to be able to generate practice exercises for a particular player based on their weaknesses, compare two players on different teams, evaluate the strengths and weaknesses of a player before signing them to a contract, determine if a player plays better on the road or at home, or determine what the best position is for a particular player. Providing effective visualizations, such as the ones included in Soccer Scoop, will lead to team managers making more informed decisions.

For data integration with our visualization, the user is required to create and load a CSV (Comma Separate Values) file based on a XLS field data template that is provided into the program before viewing either visualization.

Both the Field Viewer and Player viewer visualizations can be easily toggled, allowing for distinct player analysis of a specific team, however the Player viewer adds player-to-player comparison between teams. Such will empower a team manager to make comparative analysis regarding fine grained player characteristics at which may also assist with the coaching in the development of a tactical set play or other globally team centric decisions crucial for overall team success.

Our application possesses an intuitive design which uniquely presents all key player elements to the user. This allows for a system that can be easily adopted by even the most technically challenged team manager.

2. Background

There are many visualization applications available for the simulation of soccer games for play-by-play analysis that assist in pregame initiatives. One example of such visualization is entitled "Grass Roots Coaching" which is essentially a drill viewer visualization that encompasses the overall sequences of plays related to both offensive and defensive specific drills [11]. Other examples include "TactFOOT: Soccer Coaching Tactical Software" which allows for the creation of drills, set plays, and other training exercises that can be visualized through animations [12]. The need for coaching within the realm of tactical field dynamics

between plays is surely met with all that is currently available, however analysis tools for player-to-player comparisons is almost nonexistent and hard to find.

In this paper we present a solution for soccer player analysis; both for distinct player dynamics and player-to-player comparisons. The overall dynamics of a specific player are represented by a specific structure of data elements that were provided and used to model the requested visualization. It was hoped that three main data aspects with their respective data attributes could be used to observe the overall offensive and defensive characteristics, enabling a determination of the current skill level, and thus allowing for an overall focus onto any variables that may otherwise remain transparent; hindering any growth for player or team growth.

The following classifies the provided data structure into three main categories: execution mechanics, tactical utility, and frequent mistakes in execution. Each category is crossed by the following player elements: kick, head ball, throw-in, trap, tackling, dribbling, footwork, and ball protection. Execution mechanics for a player is comprised of precision, fluency, coordination, easiness, and rapidity. The precision attribute for all player elements are denoted by a percentage value; percentage of kick accuracy, head ball directed accuracy, throw-in accuracy to teammate, success of trap absorption, tackle success, ball position 50cm from foot while dribbling, success of feign for footwork, and success of ball protection. Fluency, coordination, easiness, and rapidity are all Likert values utilizing a scale of: unsatisfactory, satisfactory, good, and very good. Tactical utility of a distinct player is specific to all elements and each are denoted by the number of times aggregate. The value for each element denotes the following aggregate totals: number of times surprising the opposing team, head ball is used to remove ball from danger, throw-in ensures continuity of offensive attack, trapped toward the opposing goal, tackles in defending half, distance traveled in possession of ball while dribbling, ability to create a man advantage, and, number of times created advantage for opponent due to poor ball protection. For the frequency of mistakes in execution, the category comprises the player data for foot, torso, and the slow execution of movement attributes. All attributes are denoted by the Likert values as described above, however both the foot and torso denote a more fined grained characteristic as they are related to player elements. Foot attribute is broken down into the following element representations: plant foot during kick is too far from ball, pauses to jump off both legs while running during a kickball, does not keep both feet planted on throw-in, plant leg is not flexible during trap, hit opponent prior to hitting the ball during tackling, repeated use of large steps on dribbling, shifts center of gravity from one leg to another too slowly on footwork, and ball is not kept on opposite side of opponent's tackle attempt during ball protection. The torso attribute of a player represents the following element values: leaning backwards during kick and trap, moved back and forth too slowly during head ball and throw-in, does not unbalance opponent when tackling, is stiff and not leaning forward during dribbling, center of gravity too high during footwork, and not able to keep balance when physical contact made by opponent during the act of ball protection.

Conceptually, our application was designed to encompass all three categories with their respective element values all within a single viewing instance. This is visualized implementing a soccer field metaphor which utilizes a variety of unique field objects; goal and net, center field marker, ball locations, opposing team player icons, and other visual metaphors. A glyph system [8] was implemented to represent any obscure data points which could be represented naturally by field objects such as the foot and torso attributes of the mistakes in execution frequency values.

Technically, the application was developed and implemented utilizing the Java language framework. For the graphics engine, the JOGL Framework was leveraged for generating the visualizations for data representations.

3. Visualizations

3.1. Field Viewer

The Field Viewer visualization (see Figure 1) allows the manager of a soccer team to review a given player on a given team. This visualization can be toggled to show different categories. The categories are kick, head ball, throw-in, trap, tackle, dribbling, footwork, and ball protection.

The overall visualization contains smaller visualizations [9] to relay information about execution mechanics, tactical utility, frequent mistakes in execution, and slow execution of movement of a respective category.

Before explaining each visualization in detail, it is important to recognize that the entire visualization is placed on a soccer field. The key aspect to this visualization is familiarity [2]. The user needs to feel at home and must recognize the abstractions used to display the information. If the display is foreign, information might be misinterpreted.

The soccer goal represents the precision of the player for a respective category. It is appropriate that this attribute is represented with a soccer goal, as most people would associate precision with kicking a soccer ball into a goal when discussing soccer. The goal is filled up horizontally based on the precision value for the given category. The higher the value, the more filled in the goal is. The filled in portion of the goal will change color depending on the precision value. When dealing with human precognition, color is one of the first qualities of a visualization that stands out [5]. The goal is filled in red if precision percentage value is 0-25%, orange if it is 25-50%, yellow if it 50-75%, and green if it is 75-100%. It was important to fill the goal in from right to left, as a soccer team manager associates the right side of the goal in our visualization as ground level.

The center circle in the middle of the field represents the remaining attributes that make up execution mechanics, which are fluency, coordination, easiness, and rapidity. Initially, a star plot [3] was going to be used, but intersecting lines would have been confusing to the user. A similar technique was employed, as each ring represents a different attribute. The rings, from inside to out, represent rapidity, easiness, coordination, and fluency. These attributes follow a Likert scale. If a player is marked as unsatisfactory, that particular ring is red. Satisfactory means the ring will be colored orange, good means the ring will be colored yellow and very good means the ring will be colored green. Once again, color plays an important role in this visualization, as a team manager will be able to clearly identify areas of deficiency within the player's execution mechanics.

To the left of the center of the field, there is a wire-frame model of a soccer player. On this player, two regions are colored in, which are the torso and the feet. These colored regions represent the frequent mistakes in execution for the feet and torso. These attributes follow a Likert scale as well, as they determine how many times a player fails with either their feet or torso for a respective category. The coloring scheme used is the same as previously described.

The line graph with four tick marks represents the slow execution of movement attribute. A colored circle will be placed on a tick mark depending on the Likert value for a given category. Further to the right means the player is fast in execution, where as being to the far left means the player is very slow in execution. Coloring techniques mentioned previously are used here again on the circle. Coloring was added to emphasize the Likert value. If coloring was not included, labels would have been implemented beneath each tick mark to indicate to the user exactly what the value is.

Tactical utility is represented on the right side of the field with soccer balls. Tactical utility is a sum, which means the more soccer balls placed on the field, the better it is for the opponent. Choosing a soccer ball as the glyph was important [7] for a couple reasons. Firstly, using a player would have been confusing, as a player already exists on the field. Secondly, a user associates success with scoring. The user seeing a lot of balls on the right side of the field leads to the user believing that this player is outstanding despite not fully understanding the information being visualized.

The Field Viewer is a dynamic visualization, as the user can toggle between different categories (from the right side of the screen) and change players quickly (from the bottom of the screen). This will allow a team manager to quickly pinpoint deficiencies in his or her starting lineup or determine what the best pairings are for the team.

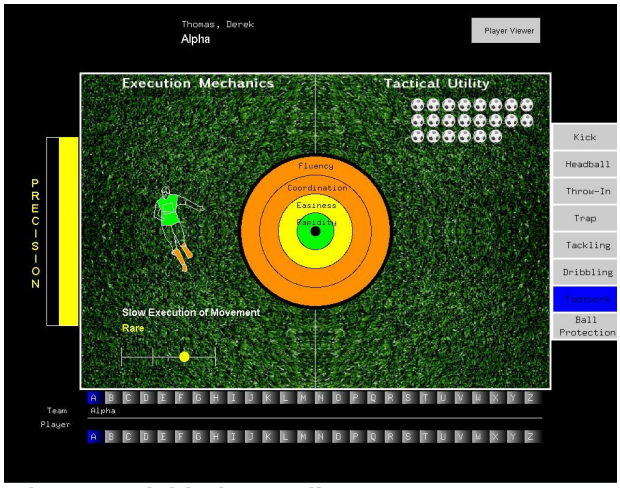

Figure 1. Field Viewer allows a team manager to view various attributes of a player dynamically.

3.2. Player Viewer

The Player Viewer visualization (see Figure 2) allows the manager of a soccer team to compare two players simultaneously. This portion of the application contains two instances of the same visualization. The visualization is a wireframe of a soccer player with a soccer ball. Each region of the player is partitioned and color coded to represent a different category. The head represents head ball, the upper legs represent kick, the arms represent throw-in, the feet represent footwork, the torso represents tackling, and the square in the torso represents trap. The color of ball represents the ability of the player to protect the ball (ball protection) and the height of the ball represents dribbling. Unlike the Field Viewer the categories that can be toggled are fluency, coordination, easiness, rapidity, foot, torso, and movement.

The same coloring techniques are used for this visualization. The Likert scale is assigned appropriate colors to each value as well as each range from 0-100 for concrete values. For the height of the ball, a similar technique was used for the visualization to represent slow execution of movement in the Field Viewer. Each tick mark is color coded so the team manager understands the meaning of the height of the ball. Labels next to each tick mark would have cluttered the visualization, so color was used instead.

The user is able to toggle their focus between the two player instances in this visualization. The user is then able to choose a different player from the selection menu at the bottom. Dynamically, the player will be updated to reflect the new player chosen. Making this visualization dynamic and the utilization of details on demand [10] allows a team manager to quickly compare players and toggle quickly between different attributes.

When developing this visualization, a key component was making it intuitive. When the user sees that the soccer player's head is red, the user should be able to deduce that they need practice at head ball. Making the right correlations between the player's

anatomy and the data is what makes this visualization effective.

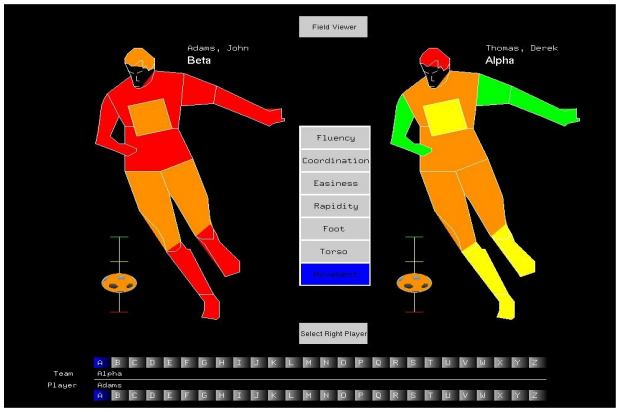

Figure 2. Player Viewer allows a team manager to compare two players dynamically and determine who is the better player based on various categories.

4. Results

The analysis of field data is critical to understanding the usefulness of Soccer Scoop system. This section explores the functionality of Soccer Scoop by outlining three specific case studies. These cases were developed and observed with the use of captured field data. Specifically, ten players of U Craiova soccer team from the Romanian soccer championship were observed over a period of five games (three home and two away). The data matrix was composed of specific fields that provided a well-rounded and very detailed synopsis of a single player in a single game. Fields such as execution mechanics, tactical utility, and frequency of mistakes were included in the data matrix.

4.1 Single player analysis using Field Viewer

Field Viewer highlights an individual player's skills and abilities. This unique visualization quickly provides all data for a single player on the screen at the same time. Through the use of familiar soccer objects (player outline, soccer ball, field, goal, etc), the data is displayed in an aesthetically pleasing manner [4]. A quick observation of the visualization can effectively point an evaluator to specific player data while still providing an overall reference point for all of the player's information. The Field Viewer provides the ability to toggle the display between each of the categories found in the data matrix (kick, footwork, dribbling, etc.).

In this analysis, as seen in Figure 3, a player was observed to have highly rated skills. Specifically, in the selected Trap category (right side of the screen), the player's information was easily displayed for quick evaluation. At first glance, a user can see that this player has a limited score of Tactical Utility by observing the small number of soccer balls in the upper right corner of the screen. Another glance will show the torso of the player highlighted in yellow while the feet are red. This indicates the level of frequent mistakes that this player has made in torso field skills and footwork. A final glance at the visualization will show the specific areas that this player excels in. The goal is filled and green, thus showing this player's precision advantage. Each ring in the center of the field is highlighted green to show that this player has strengths in fluency, coordination, easiness, and rapidity. According to the visualization, this player possesses more strengths than weaknesses, thus representing the quality that this player brings to the team.

Figure 3. Field Viewer for a player ranking above average in the Trap category.

4.2 Comparison of two players using Player Viewer

The ability to compare players during a game or during practice sessions can be a challenge. Team managers from various sports can observe individual performances and take notes, but the focus tends to be on a single player and does not allow a team manager to quickly identify the differences between players. For the best player-to-player comparisons, Player Viewer will be observed for this case. Once the visualization is launched and the two players are selected, the user is presented with a color coded comparison of the players. This color coding of the players makes their abilities immediately apparent to the user. The user can chose between different categories (fluency, easiness, rapidity, etc.). When a category is selected the visualization will be redrawn to display the data from the corresponding column in the data matrix.

As an example, looking at Figure 4, the user has selected the easiness category in Player Viewer. When the visualization is drawn it can easily be inferred that the player on the left performs throw-ins with greater ease than the player on the right. This is determined by the color of the player's arms. The player on the right has arms that are highlighted as red while the player on the left has arms that are highlighted as yellow. As mentioned above, the soccer ball next to each player

encodes two pieces of information: ball protection and dribbling. This visualization reveals to the user that although the player on the left performs dribbling with more ease than the player on the right (represented by height of the ball), the player on the right can more easily protect the ball as shown by the green color of the ball.

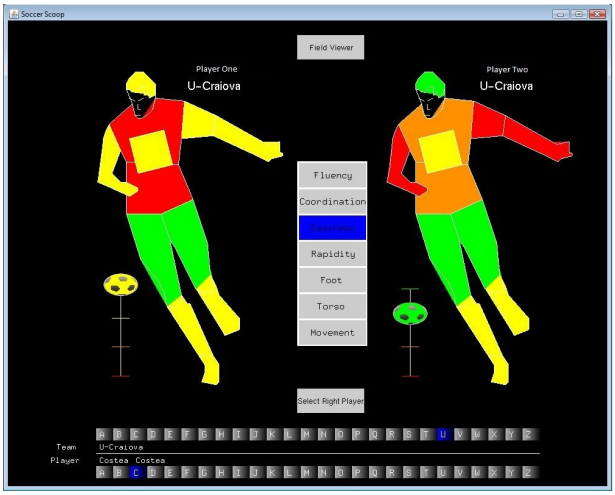

Figure 4. Player Viewer after real data has been entered for two different players. The Easiness category has been selected.

4.3 Comparison of the entire team analysis (average) between home and away games

Team managers seem to be limited to only evaluating a player in the setting/location at hand since the player cannot be in two places (home and away) at the same time. This complication forces the team manager to rely on visual or statistical memory of the player's performance in a different setting. Even with the separate captured data from home field and an away field, accurate estimates of a team's performance may be unclear due to the fact that comparing data matrices side-by-side is not easy to evaluate at a glance [6]. Soccer Scoop provides a better way to give team managers insight into how their team performs at home and on the road. Since sport teams have demanding schedules during a single season, the overall team performance at various locations deserve careful evaluation. Critical data captured over a period of games performed at a team's home field and during away games could determine key strengths and/or weaknesses of a team. This particular case attempts to show a glimpse into what Soccer Scoop can be used for in determining the performance of the team as a whole during home games and away games. Captured field data was consolidated into two categories. The first category grouped all player data for home games. The second brought together all data for away games. In each of these two categories, all player data was averaged together for all home games and all away games. In essence, this case sought to display the entire team average for home games in a single player visualization. Similarly, a single player visualization was used to represent the overall team average for away games. This separation of home and away game data allowed Soccer Scoop to treat each data set each as a single player, thus being able to display these player visualizations side by side for evaluation. The visualization generated from this data can be seen in Figure 5. At first glance, this particular comparison shows that the team generally performs similarly in rapidity regardless of playing at a home or away game. Careful evaluation further reveals that the rapidity of the team's trapping and dribbling is better in away games. It can also be seen that the team is better at protecting the ball at home. These discoveries can assist the coach in developing exercises and practice sessions for the areas that need improvement.

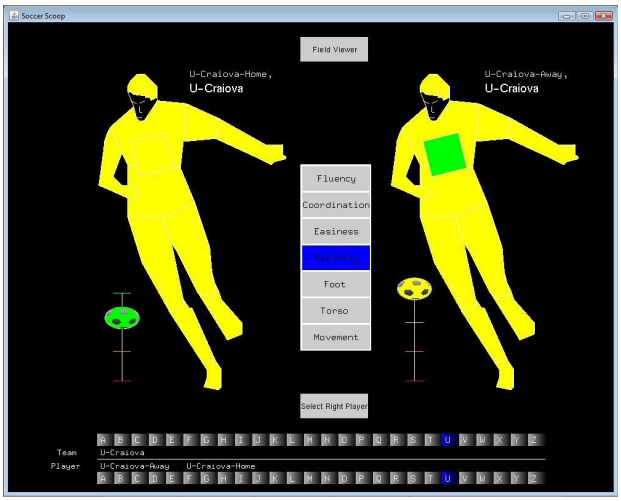

Figure 5. Comparison between a team's average performance at home versus average performance at away games.

4.4 Results Summary

The case studies described above represent a small subset of possibilities that Soccer Scoop can be used to evaluate. The focus of the cases was on the performance of a single player using the Field Viewer, the comparison of different players using the Player Viewer, and the evaluation of the data averaged from the entire team performance during home games versus away games. The intent of the Soccer Scoop tool is to empower coaches with the ability to analyze and evaluate the team efficiently and effectively. With an intuitive representation of the player's skills highlighted by a basic color scheme, decisions can be made and practice exercises can be determined for the team and individual players. Soccer Scoop quickly displays not only the best skills of each player, but also shows the areas that need improvement. Ultimately, Soccer Scoop can be used by coaches to strengthen the overall performance of their team.

5. Future Enhancements

In this section, future enhancements to Soccer Scoop are considered. These changes would increase the flexibility and control of the program as well as provide additional analysis control to the program. Firstly, a possible future improvement of Soccer Scoop will ultimately include a more streamlined interface with more emphasis on data control than is currently implemented. For example, a feature may be included that would allow for additional data imports that would append addition data to the currently available data. Additionally, more customization and personalization options may be implemented and made available to users that would allow for a quick launch of previous application sessions; allowing for convenient analysis and to avoid re-import frustrations as posed by the re-launch of a new session.

6. Conclusion

We present Soccer Scoop, a visualization application that is comprised of two main soccer player visualizations. Both represent distinct player characteristics as related to specific player dynamics, however it is also possible to establish an analysis of player-to-player comparisons in the same team or between two teams, thus fulfilling the need for coaching analysis related to player specific skill-set data. Our quick-glance observation technique is made possible by leveraging a soccer field metaphor that encapsulates a complete set of skill-set categories with associated attributes and cross references each element of a player represented by natural objects located on the field. Ultimately, through quick observation and analysis as implemented by our application, coaches will benefit from the use of Soccer Scoop. Specifically, coaches will achieve the ability to categorize players, focus in on a skill quality or deficiency, and perhaps establish better strategies in terms of overall team development and success.

References

[1] Jack Ox. Visualization and the art of metaphor. Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition, pages 307-308, 2007.

[2] Phil Turner. Being-with: A study of familiarity. Interacting with Computer, Volume 20, Issue 4-5, pages 447-454, 2008.

[3] Michael Friendly and Ernest Kwan. Effect ordering for data displays. Computational Statistics & Data Analysis, Volume 43, Issue 4, pages 509-539, 2003.

[4] Youn-kyung Lim, Erik Stolterman, Heekyoung Jung, and Justin Donaldson. Interaction gestalt and the design of aesthetic interactions. Proceedings of the 2007 conference on Designing pleasurable products and interfaces, pages 239-254, 2007.

[5] P. K. Robertson and M. A. Hutchins. An Approach to Intelligent Design of Color Visualizations. Scientific Visualization, Overviews, Methodologies, and Techniques, pp. 179-190, 1994.

[6] A. Telea, P. de Hillerin, and V. Valeanu. Visualization of multivariate athlete performance data. The Seventh IASTED International Conference on Visualization, Imaging and Image Processing, pages 123-128, 2007.

[7] B. Yost and C. North. Single complex glyphs versus multiple simple glyphs. Conference on Human Factors in Computing Systems, pp. 1889-1892, 2005.

[8] Matthew O. Ward. A taxonomy of glyph placement strategies for multidimensional data visualization. Information Visualization, Volume 1, Issue 3-4, pages 194-210, 2002.

[9] Eliane R.A. Valiati, Marcelo S. Pimenta, and Carla M.D.S. Freitas. A taxonomy of tasks for guiding the evaluation of multidimensional visualizations. Proceedings of the 2006 AVI workshop on Beyond time and errors: novel evaluation methods for information visualization, pages 1-6, 2006.

[10] Darius Pfitzner, Vaughan Hobbs, and David Powers. A unified taxonomic framework for information visualization. Proceedings of the Asia-Pacific symposium on Information visualization, Volume 24, pages 57-66, 2003.

[11] Soccer & Football Coaching, Drills, & Training in the UK, Europe, & America. Retrieved from http://www.grassrootscoaching.com, 2005.

[12] TactFoot: Soccer Coaching Tactical Software. Retrieved from http://www.tactfoot.com, 2010.

[13] Miller, Todd and Stasko, John, "The InfoCanvas: Information Conveyance through Personalized, Expressive Art", Short Paper, CHI 2001 Extended Abstracts, Seattle, WA, April 2001, pp. 305-306.

[14] Niklas Elmqvist, John Stasko and Philippas Tsigas, "DataMeadow: A Visual Canvas for Analysis of Large-Scale Multivariate Data", Information Visualization, Vol. 7, No. 1, Spring 2008, pp. 18-33.