



From Sports to Science: Using Basketball Analytics to Broaden the Appeal of Math and Science Among Youth

John F. Drazan, Amy K. Loya, Benjamin D. Horne, and Ron Eglash
Rensselaer Polytechnic Institute
Email: john.drazan@gmail.com

Abstract

The fusion of sports and analytics has not only revolutionized the way professional basketball is played, coached, and managed; it also has the potential to revolutionize the way we educate the next generation. Sports analytics can provide a rigorous, yet tangible application of math and statistics in which youth perform data gathering and analysis directly linked to their own improved on-court performance. This allows students who are not traditionally engaged in STEM (Science, Technology, Engineering, and Mathematics) to be intrinsically motivated to use STEM concepts as a tool for basketball training. We hypothesized that once STEM has shown applicability towards improving their own basketball skills, it would also become more attractive as a career path. Here we provide preliminary data consistent with that hypothesis that shows that these clinics increase athletes self-perception of knowledge and interest in STEM. This approach can be integrated within the sports organization's existing youth programs to promote STEM education as a public good, and can be replicated in other educational communities, both formal and informal.

1. Introduction

As the application of science to sports has gained mainstream acceptance, researchers have expanded the search for new applications. Our work shows that STEM (Science, Technology, Engineering and Mathematics) education, including that of youth populations currently underrepresented in the STEM workforce, can greatly benefit from the incorporation of sports science, specifically sports analytics, as a component of youth programs. Black, Latino and Native American students, as well as white and Asian students from low-income backgrounds, are all underrepresented in STEM (Landivar, 2013; Lee, 2015; Rothstein, 2013); yet their interest in sports is often quite profound. Sports science can provide a bridge between these two worlds, motivating and guiding underrepresented students (Drazan, Cooke, & Eglash, 2016). However, this bridge must be properly constructed. For example, calculating the parabolic motion of a basketball shot is a shallow connection. Sports analytics, specifically in the context of physical practice, may offer a deeper interaction in which youth are intrinsically motivated by their own interest in athletic improvement. This paper offers preliminary evidence for this hypothesis.

For students who are not underrepresented in STEM, their passion for gadgets, rock collecting, or telescopes is well served by after school clubs, and translates directly into career paths. The role of out-of-school STEM activities has been well defined as beneficial for increasing student engagement, scholastic achievement, and ownership of STEM topics. (Bell, Lewenstein, Shouse, & Feder, 2009; Gottfried & Williams, 2013) For example, students who participate in robotics clubs have a high

likelihood of pursuing a lucrative career in the STEM fields. One review of graduates from a major high school robotics program reported that that 78% of their participants subsequently took postsecondary STEM classes and 60% secured at least one science or technology-related internship. (Melchior, Cohen, Cutter, & Leavitt, 2005). However this may be due to the fact that students who already have STEM interest are mostly likely to participate, while students from low income backgrounds are less likely to participate (Gottfried & Williams, 2013). Thus, the students in most need of STEM enrichment are the ones who are least likely to sign up for additional STEM activities (President, 2010). As one study centered on informal STEM outreach using Lego robotics stated (Karp & Schneider, 2011) :

Voluntary extracurricular programs tend to have a strong self-selection of students. It is quite likely that only students with an interest in robotics and science in the first place will choose to participate in the program.

To address this “interest-gap” it is necessary to find a method to bring STEM into out-of-school activities that are accessible to all students regardless of preexisting interest.

Participation in sports is far more popular among underrepresented students, but the career trajectory within sports is extremely difficult. Out of the almost 1 million male and female high school athletes, only 3.6% of these students play basketball in college, and only 8.7% of these college athletes continue to play some form of professional basketball. An average high school athlete only has a 0.3% chance to play professionally (NCAA Research, 2016). This analysis excludes students who simply enjoy playing basketball or following professional sports in their free time, but don’t play for their high school team. Thus, while participation in sports brings many laudable benefits such as learning teamwork, dedication, and cooperation, it does not translate directly to a career path as seen in out of school STEM activities. The fusion of youth sports programs with sports analytics is an excellent opportunity to serve as an introduction to STEM within the arena of sports to broaden the appeal of STEM to underrepresented youth.

Improvements to the quality and quantity of available data in basketball has led to an analytical revolution within the NBA. Player tracking data, generated by companies such as SportsVU, has been used by researchers to develop analytical tools that provide insight into offensive and defensive efficiencies of NBA players and teams. These metrics provide players with information on weak spots in their game, how to guard opponents, and how to optimize on-court performance. A huge amount of resources has been poured into analytics departments around the NBA so that teams can leverage this knowledge as a competitive advantage. As the data sets and analytics have become more complex, it has become necessary to develop ways to easily interpret this information. Hence, advances in analytics have been coupled with advances in graphical representations of complex data sets. An excellent example of this are the shooting and defensive efficiency maps created by Kirk Goldsberry (Franks, Miller, Bornn, & Goldsberry, 2015; Goldsberry, 2012). These accessible graphical representations of data have allowed analytics to penetrate the media space. Game highlights often include heat maps and there are entire popular websites dedicated to the statistical analysis of sports. Unfortunately, while the public’s exposure to data analytics in sports is at an all time high, the high cost of tracking systems and the expertise required for this analysis have created a barrier to the

widespread use of these techniques outside the NBA. This means that the interest that youth athletes or coaches may have in using analytics as a tool to understand their own performance has no outlet.

Fortunately, the recent advances in basketball analytics are built upon a legacy of statistical analysis of basketball that was developed before the advent of these advanced tracking technologies and analytics. Valuable performance metrics such as shooting percentage and scoring efficiency can be determined by collecting data with a pen and paper and calculated using simple arithmetic. We can provide youth players and coaches with an important training tool by combining these accessible analytical basketball metrics with modern graphical representations of data such as heat maps. These heat maps can show players where they need to improve as shooters, where they should be shooting from, and how they can be effective players. In effect, many of the benefits of modern basketball analytics are being expanded to the general population who otherwise don't have the technical resources or expertise to participate in the modern analytics revolution.

We believe that the benefits of this expanded reach of analytics are twofold; it not only allows players and coaches to use analytics to improve their own performance, but it also provides an avenue for youth to experience scientific inquiry within a venue at which they are intrinsically motivated to improve. This increases the impact of youth sports programs: not only are students benefiting from the values imparted by sports participation, but basketball is also being used as a venue for informal STEM learning. To investigate the potential of shooting analytics as a method to increase youth athletes' understanding of basketball performance and training through the use of STEM topics, we created an open source shooting analytics program that generates personalized heat maps based on user collected data. This approach was designed to be scalable through its use of open source software so that it can be used by other youth basketball programs. We used this program as part of shooting clinics performed at local youth programs and evaluated changes in student attitudes towards basketball training and STEM within sports analytics.

2. Methods

2.1. Overview of Shooting Clinic

To achieve this aim, we developed an open source shooting analysis program for use at basketball clinics and practices. We recruited local youth coaches to host these programs as part of regularly scheduled clinics and practices. The program was seen as a shooting clinic that would assist in player development and coaching rather than strictly as a STEM enrichment activity. Before the start of practice, all players completed a survey, which served as a baseline measurement of their existing attitudes toward basketball training and STEM, respectively.

After a brief introduction to the use of analytics within basketball and the importance of data collection, players participated in their routine warm-up. Then players engaged in the shooting drill that allowed us to collect data to create personalized heatmaps. Players divided into groups of 4 to 5 per basket and were given a data entry sheet to record shooting results. The players followed a predetermined shooting pattern, taking a total of 10 shots at 14 locations, that varied in angle and proximity to the basket. The players rotated periodically throughout the drill to alternate rebounding, passing, shooting, and data entry roles. This helped to prevent shooter fatigue and it allowed each student to experience data collection.

Once all players had completed the drill and submitted their results, the data was entered into the shooting analytics program. As the individual heat maps were being constructed, sample heat maps were given to each player to guide a discussion about how these maps can be beneficial for informed training, game time decision making, and defensive assignments. Following this analysis, players received their own personalized heat maps, and coaches were given access to the heat maps for each of their players.

2.2. Personalized Heat Maps

The raw data for the number of made baskets per location was collected and entered into the computer program for heat map generation. Python programming language and built-in matplotlib toolbox was used to construct two types of heat maps, as shown in Figure 1. The first calculates shooting percentage at each location and displays the results for each shot location with color indicating the percentage. This percentage graphic provides players with a visual representation of where their strengths and weaknesses are as a shooter. For example, this information can be used by coaches or players to identify and address specific limitations in a player's skill set with respect to his or her position.

The second heat map uses the shooting percentage at each location to calculate a weighted value per its point worth in a game (i.e. 2 points vs. 3 points). This results in an efficiency metric at each location, with the color indicating a value between 0 and 3 points per shot. In addition to providing players with information regarding their own performance, the efficiency heatmap enables discussion of in-game offensive and defensive strategies. Alternatively, the resultant heat map depicts locations from which the player is most effective and therefore should take more shots. This shooting analytic program is open source and is freely available online at <https://community.csdt.rpi.edu/cms/application-contexts/diy-sports-science-lab/>.

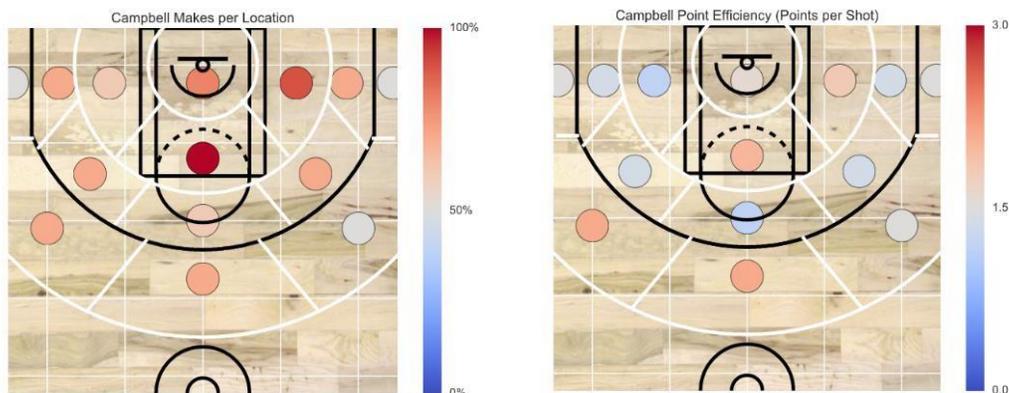


Figure 1: Example heat maps. The left shows makes per location, the right shows point efficiency.

2.3. Evaluation of Participants

The survey used for the program evaluation was administered both prior to the shooting clinic and following the conclusion of the discussion on heat map applications. Program participants reported demographic information, including what school they attended, in addition to their responses to

questions that measured the participants' attitudes towards basketball training, STEM, and applications of STEM within basketball.

It has been shown that an individual's self-efficacy and the self-perception of their ability can impact their performance within that activity (Bandura, 1982). Therefore, this enhancement of an athlete's self-efficacy regarding basketball training may serve to improve the athlete's capacity for and quality of training. Student attitudes towards science topics and careers are an important determinant in engagement in STEM learning and activities. By changing the attitudes of student athletes towards STEM, they can be more easily engaged in future STEM activities in the classroom (Germann, 1988). We used 15 likert style questions (5 point scale), as well as three short answer questions, to measure participant attitudes. The three intended outcomes for our program are as follows:

1. Contrary to previous STEM outreach efforts, the participants in this activity would not be self-selected for STEM enrichment.
2. Participation in the collection and analysis of shooting data will increase participants selfefficacy in basketball training
3. The application of STEM within basketball using shooting analytics will increase students awareness and interest in the STEM fields.

In effect, shooting analytics would provide a bridge between basketball and STEM for students who are not typically exposed to STEM topics outside of school.

2.4. Statistical Analysis of Data

To test the statistical significance of our results, we used Wilcoxon signed-rank test to compare the participants' responses between the pre- and post-tests. Wilcoxon sign-rank test is a non-parametric hypothesis test that can be done on ordinal scale data to show a significant difference between two distributions. In this work, we used Wilcoxon sign-rank test to show that the treatment of our shooting program makes a significant shift in the distribution of survey answers for each question. We opted to use Wilcox rather than a one-way ANOVA test or a paired T-test because we cannot assume our data is normally distributed or that it has an interval scale (Allen & Seaman, 2007). Along with question by question analysis, we grouped questions into two categories to test the overarching concepts outlined in hypotheses 2 and 3: basketball self-efficacy and STEM interest, respectively. As these categories intersect at the use of shooting analytics within basketball, several questions were used in both categories. To create each group, we took an average of pre- and post-test answers, independently. To demonstrate this technique's validity, we used Chronbach's alpha to show internal consistency within each of the categories. Chronbach's alpha is typically used as a reliability measure of a psychometric test. This test is done by computing a correlation score on a set of tested items that reflect a common construct. In this study, we split our survey questions into the two concept categories, treating each category of questions as its own psychometric test, to calculate the reliability of each construct set of questions. If both concept categories are internally consistent, we can use this combined statistic to illustrate the overall distribution shift for a given concept and test for significance (Gliem & Gliem, 2003). It is commonly held that Cronbach's alpha values of between 0.7 and 0.8 acceptable internal consistency and values between 0.8 and 0.9 good internal consistency (George & Mallery, 2003).

3. Results

3.1. Outcome 1: Contrary to previous STEM outreach efforts, the participants in this activity would not be self-selected for STEM enrichment.

We performed the program on six different occasions at basketball clinics and practices. We collected a total of 98 matched pre- and post-tests. As seen in Table 1, the program participants were a diverse cross sample of students. 35% of participants were female and 53% of participants identified as either African American or Latino. 62% of students attended urban high school districts that serve over 50% economically disadvantaged students. At the start of the clinic, over 50% of students reported that they did not enjoy using STEM outside of school. In contrast, over 70% of students reported that they wanted to play basketball in college. These responses and student demographics support our contention that using analytics within basketball provides STEM outreach to a population of youth who are not predisposed for STEM engagement. As self-selection is a persistent problem with reaching new students with STEM outreach efforts, this broadened engagement is already a significant outcome.

Category.	#Students
Female	44
Male	54

(a) Gender

Category.	#Students
Urban	61
Suburban	37

(b) School Type

Category.	#Students
African American	38
White	44
Asian	2
Hispanic	8
Native American	1
Pacific Islander	1
Multiracial	4

(c) Ethnicity

Category.	#Students
7th Grade	11
8th Grade	8
9th Grade	22
10th Grade	29
11th Grade	18
12th Grade	10

(d) Grade

Table 1: Participant Demographic Data

3.2. Outcome 2: Participation in the collection and analysis of shooting data will increase participant's self-efficacy in basketball training

We found that the inclusion of basketball analytics as a training tool increased participants' self-efficacy with regards to basketball training. Of the Likert questions identified as training related, only one of them, Question 9, did not exhibit a statistically significant increase between the pre- and post-test as shown in Table 2.



Survey Results: Training Concept		
QID	Survey Question	P-Value
1	How confident are you in your ability to evaluate your strengths and weaknesses as a shooter?	$3.4942e - 06^*$
3	Do you know where on the court you need to practice shooting from?	$1.8271e - 04^*$
5	Do you think you know what you need to practice in order to improve your shooting skills?	$1.6311e - 06^*$
6	How familiar are you with sports analytics (the use of math and statistics to measure player performance in sports)?	$1.1024e - 05^*$
8	Do you think that you know how to use sports analytics to improve your own shooting abilities?	$5.7112e - 10^*$
9	How interested are you in learning about sports analytics to improve your shooting abilities?	$4.8986e - 01$
10	Do you want to play basketball in college?	$2.6207e - 02^*$
11	Do you think that you know enough about basketball to reach your goals as a player?	$3.7114e - 04^*$
12	Are you interested in learning about how you can use math and science to become a better basketball player?	$1.8623e - 02^*$

Table 2: Survey Wilcoxon sign-rank results for Training Concept questions. P-values for statistically significant pre/post differences ($p < 0.05$) indicated by “*”.

Participant perceptions of their own ability to understand their own performance in basketball as measured by shifts in responses between the pre- and post-tests. Participants reported a greater confidence in their knowledge of their own strengths and weaknesses (Q1), an understanding of their own shooting performance with respect to court location (Q3), their own knowledge of practice (Q5), and their general expertise as basketball players (Q11). Participants began with a relatively high self-evaluation of their knowledge on average; however, participation in the shooting analytics clinic dramatically reduced the number of low scores for all questions, as seen in Figure 2. Similarly, participants’ answers to long answer questions shifted as well. Several responses shifted from generalized, “I need to work on my footwork and form” to including spatial information such as, “I need to work on mid range and 3 pointers.” The shifts indicate that the use of heat maps within a youth sports practice increases participants’ self-perception of their own expertise within basketball.

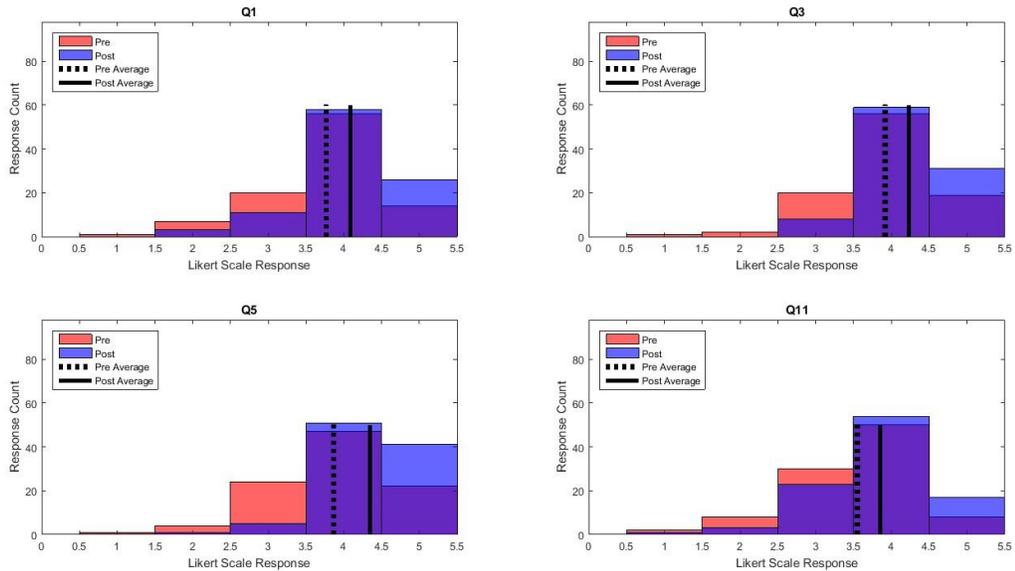


Figure 2: The shift in participant responses on questions regarding basketball training.

We also observed a shift in three out of four responses related to the role of sports analytics within basketball. Participants reported increased familiarity with sports analytics (Q6), an increase in an ability to use sports analytics to improve their own training (Q8), and an increased interest in applying math and science to improve at basketball (Q12). Interestingly there was minimal change in the distribution of responses regarding a student's interest in learning how to use sports analytics to improve at basketball (Q9). When comparing the initial responses for questions 6, 8, and 12 to that of 9, 9 had a higher incidence of more positive responses. Question 15 asked participants if they thought that sports analytics were important to professional basketball players and franchises. Over 85% of participants indicated that analytics was either important or very important on the pretest. As such, student athletes would already be very interested in using analytics within their own games, thus reducing the amount that this metric could be improved. Regardless, these results indicate that the connection between sports and analytics was strengthened due to participation in this program.

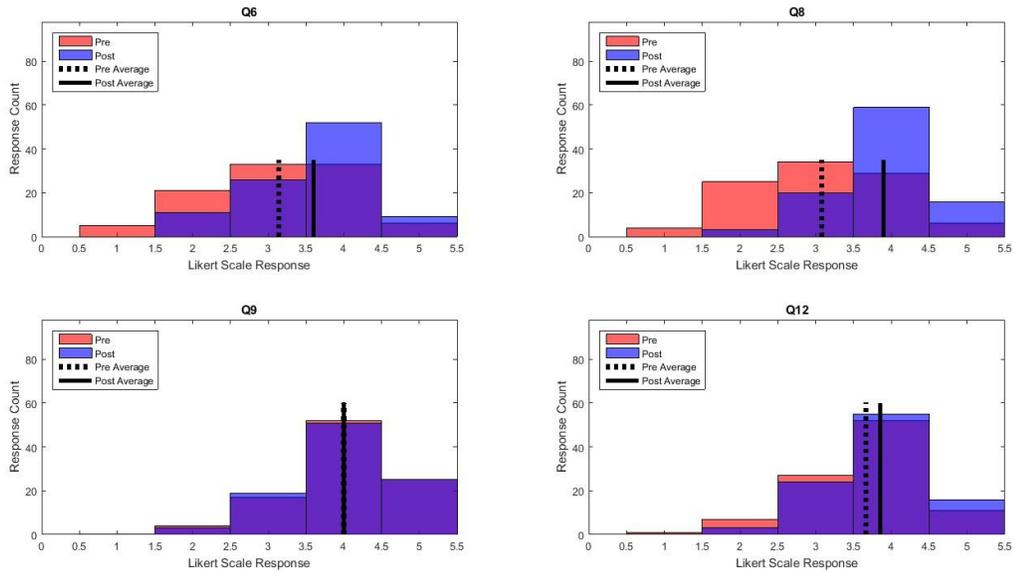


Figure 3: The shift in participant responses on questions regarding analytics as a training tool.

We observed a statistically significant positive shift in the overall metric that we developed to characterize the self-efficacy of student athletes to improve training through the use of sports analytics increase with a p value of $5.8187e-13$ (Figure 4). The Cronbach alpha value of this combined metric was 0.800. This indicates good internal consistency within the questions used to evaluate student-athlete perception for using sports analytics for training. These results highlight the potential of heatmaps to improve training among youth because self-efficacy is a crucial component for success in an activity.

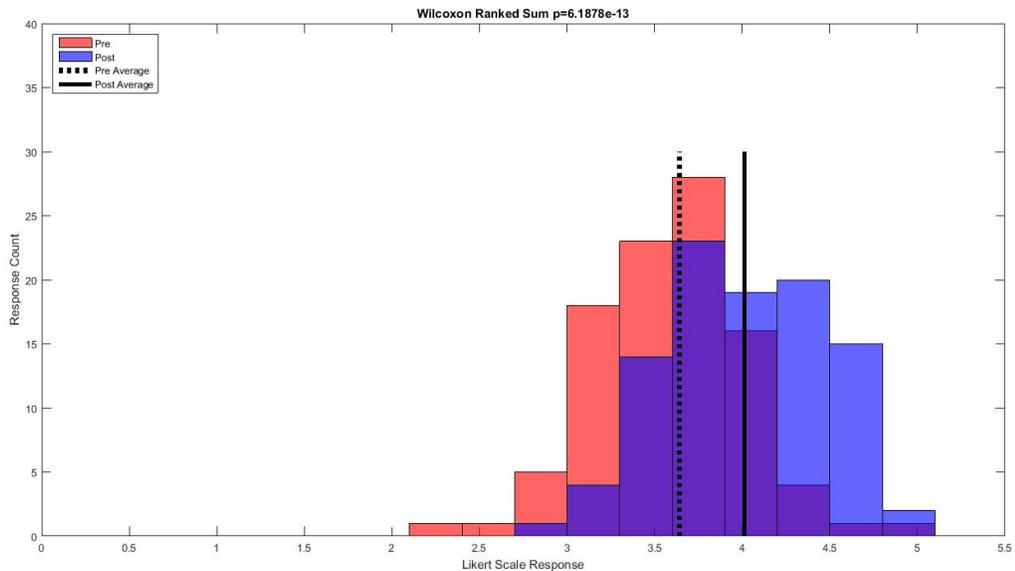


Figure 4: The shift in the combined metric that reflects participant connections between analytics and basketball training.

3.3. Outcome 3: The application of STEM within basketball using shooting analytics will increase students awareness and interest in the STEM fields.

As we observed a statistically significant increase in the metric indicating participant self-efficacy regarding sports analytics as a training tool, we can now analyze the responses regarding STEM within analytics to determine if we also observe an increase in general student-athlete interest in STEM. Of the 10 questions included within this metric, seven out of ten had statistically significant increases as shown in Table 3.

Survey Results: STEM Concept		
QID	Survey Question	P-Value
2	How important is it to you to do well in your math and science classes?	$9.6531e - 01$
4	Do you think that the topics that you learn in math and science class can make you a better basketball player?	$5.5943e - 10^*$
6	How familiar are you with sports analytics (the use of math and statistics to measure player performance in sports)?	$1.1024e - 05^*$
7	How interested are you in studying science, math or engineering in college?	$8.9011e - 05^*$
8	Do you think that you know how to use sports analytics to improve your own shooting abilities?	$5.7112e - 10^*$
9	How interested are you in learning about sports analytics to improve your shooting abilities?	$4.8986e - 01$
12	Are you interested in learning about how you can use math and science to become a better basketball player?	$1.8623e - 02^*$
13	Do you enjoy using math and science when you are outside of school?	$1.0878e - 04^*$
14	Do you think that what you learn in math and science class can be used in sports analytics?	$1.2415e - 03^*$
15	How important do you think sports analytics are to professional basketball players and teams?	$7.0311e - 02$

Table 3: Survey Wilcoxon sign-rank results for Stem Concept questions. P-values for statistically significant pre/post differences ($p < 0.05$) indicated by '*'.

Examining individual responses centered on STEM engagement, we observed a significant positive shift in participant responses regarding the belief that topics in math and science can make someone a better basketball player (Q4), interest in studying STEM topics in college (Q7), enjoyment in using STEM outside of school (Q13), and the belief that STEM topics are applicable to sports analytics (Q14), as seen in figure 5. The shifts between pre- and post-tests for Q4 and Q13 were especially pronounced with sharp reduction in negative responses. Additionally, most of the students who reported the lowest interest in for Q7 increased leaving only 4 respondents reporting that they were “not interested at all” in STEM in college. These trends were captured by one student's response to the question, “Do you think that you can use what you learn in math class to improve your performance in basketball?” The student's initial response was “No, because we are learning about proofs and

proofs don't help you with basketball." Their final response was, "Yes, because you need to know percentages to know where you can shoot from." This student entered the clinic seeing STEM as an academic topic that had no impact on basketball and left the clinic with a concrete way to apply STEM to improve basketball performance through the use of analytics.

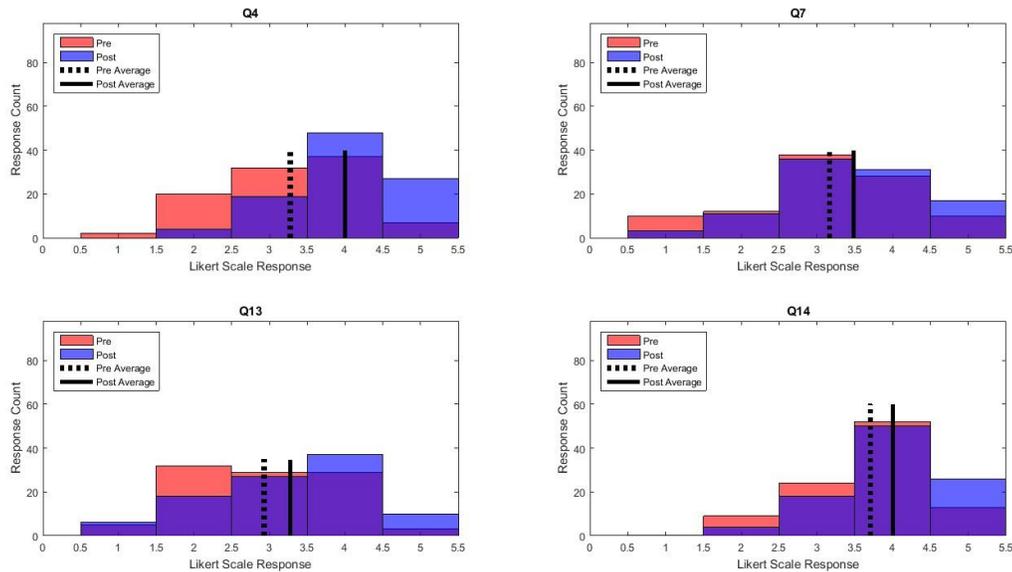


Figure 5: The shift in participant responses on questions regarding STEM interest and its application to basketball.

When we combined the responses that formed the overall metric for STEM interest we observed a statistically significant shift between the pre-test and the post-test as shown in Figure 6. Using the Wilcoxon rank sum test we found a P value of $1.4326e-10$. The Cronbach alpha value of this combined metric was 0.816. This indicates good internal consistency within the questions used to evaluate STEM interest in participants. In the pre-test we observed a small group of participants who scored very low on this metric as an outlier and after the treatment these students rejoined the overall group. The observed shift indicates that participants became more interested in STEM and applying STEM to basketball due to participation in the treatment.

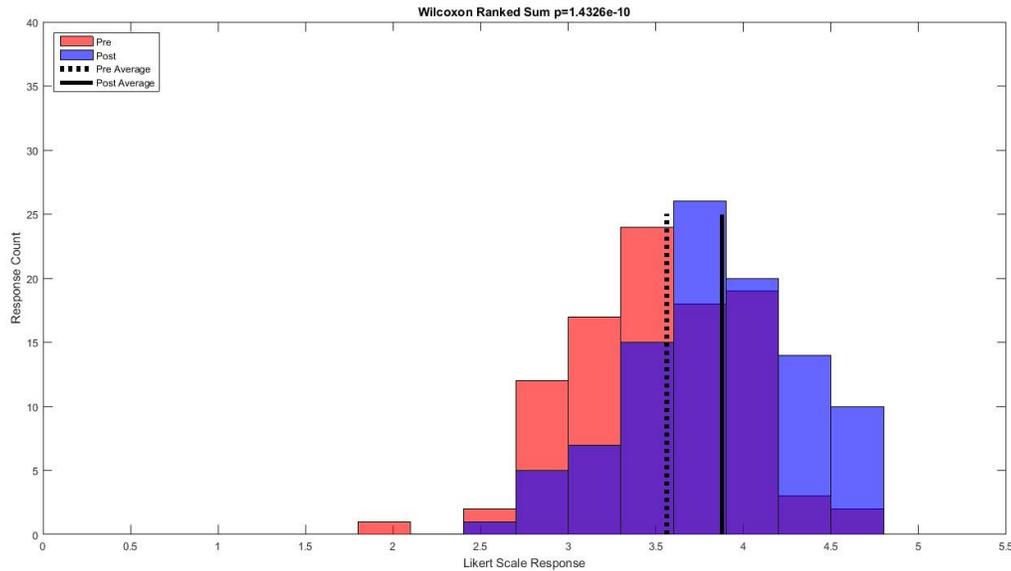


Figure 6: The shift in the combined metric that reflects participant attitudes and interest in STEM and analytics.

In summary, our results supported the targeted outcomes that we had in mind in when designing the clinic. The demographics and reported interests of the participants were atypical for out-of-school enrichment activities and indicate that we reached a new section of students with our program. This is important because it addresses the issue of self-selection for STEM enrichment activities. Participants reported increased knowledge of their own performance and ways to address these limitations. The good intergroup correlation coefficient between basketball training questions and sports analytics questions provides support to the idea that this was achieved through a new understanding of basketball through analytics. As self-efficacy is an important component of future success, these results indicate that the use of basketball analytics communicated through heat maps is a promising avenue for improvement for youth athletes.

We also observed an increase in awareness of the applications of academic STEM subjects within basketball. We saw a similar increase in interest in STEM topics across participants. The good intergroup correlation coefficient among questions regarding connections between STEM and sports analytics provides evidence that these two outcomes are correlated. Motivation and interest is a key ingredient for STEM success and the use of basketball analytics within youth programs has potential to engender these characteristics in previously unengaged youth.

While these results are promising, this study is the just the first step in an exciting new research direction. The development of interest in an activity can be broken into four steps: the first is situational interest, characterized by short term, externally motivated interest. The second is maintained situational interest, characterized by persistent engagement within an activity. The third is emerging interest in which the participants begin to realize the relevance of the activity to their

identity. The final stage is individual interest in which an individual strongly identifies with an activity and is intrinsically motivated to improve within that domain (Maltese, Melki, & Wiebke, 2014).

While students cannot be expected to transverse this entire process through participation in an hourlong shooting analytics clinic, the results of this study provide preliminary evidence that the use of shooting analytics within youth basketball is a promising avenue for STEM engagement. It is possible that additional STEM topics can be linked to such experiences, providing longer term exposure over the course of formal education. And even short term activities such as those described in this paper may play an important role in sparking initial interest for entering the STEM pipeline. Additionally, improved evaluation methods for scholastic performance metrics as well as athletic performance metrics following this program would serve to validate the importance of short term interest and attitude assessments using likert scale type questions.

4. Discussion

By incorporating analytical techniques such as data collection, analysis, and interpretation within the context of basketball practice, the students that we work with are not self-selected for STEM enrichment. The required interest for participation is centered on basketball, and thus, more conducive to a population of underrepresented students that STEM outreach is not typically serving. Persistence within the STEM fields has been associated with self-identification as scientists and intrinsic motivation for learning its subject matter. We have provided evidence that by expanding the sense of ownership and expertise that a youth feels for basketball into STEM topics using data collection and heatmaps, traditionally marginalized students may be authentically introduced to the STEM disciplines.

During discussions with groups of players after data collection, players began to identify trends within their own and others' performances. We observed coaches using these graphics to have discussions with players about the importance of shot selection. Interestingly, some of those players then used their heat maps to respond to these critiques. Memorably, one coach and player came to the conclusion, "can't argue with that data" when discussing the value of a power forward shooting the corner three. During one clinic, players complained that taking 140 shots in one session introduced fatigue as "experimental error" to the collected data. Basketball provides a venue for the students and coaches to experiment with the scientific knowledge that they learned, in effect turning them from consumers to producers of scientific thought. By scaffolding scientific knowledge within their own knowledge of basketball, the expertise that they associate with basketball begins to be transferred to the STEM disciplines (Drazan et al., 2016; Eglash & Garvey, 2014).

By introducing analytics as a tool for player improvement, STEM knowledge can be assimilated into the identity of successful basketball players. Just as students in robotics clubs use STEM to compete in robotics competitions, the use of analytics within youth sports has potential to allow athletes to learn technical concepts to win basketball games. This will allow them to think of basketball analytics as something that enhances basketball training and performance, thus, making it analogous to lifting weights or spending time in the gym. These basketball training activities are seen as important to the development of a player and are held in high esteem by peers and coaches. If STEM based analysis of basketball can be seen as a similarly laudable activity, youth will be intrinsically motivated to

understand these topics and apply them to their own training. As the communication of analytical information between statistical staff, coaches, and players is sometimes a challenge, the introduction of analytics early in a player's development may make these techniques more accessible to all parties. Just as strength training and shooting drills are an important part of player development, an understanding of analytics may one day be just another tool for athletic success.

This study shows that the analytical revolution that has swept the NBA has potential impacts outside of the competition and championships. We have shown preliminary data that the use of analytics in youth programs has the potential to provide youth with motivation to succeed on and off the court. The advantage of the approach outlined in this study is that it can be integrated into existing youth basketball programs. As many NBA franchises already spend considerable resources on both youth development programs in the form of clinics and camps as well as employ many data analytics staff, it is simple to incorporate both efforts to enhance the beneficial impact of development for youth participants as athletes and as students. The use of analytics within youth programs is a promising way to capitalize on the passion that youth have for basketball to promote STEM education.

5. Conclusion

This study provides preliminary evidence regarding the efficacy of using basketball analytics as a method to increase youth interest in the STEM fields while improving basketball training. While the importance of analytics and STEM careers within sports has increased dramatically in recent years, these techniques have yet to impact youth basketball programs on a wide scale. The application of STEM within sports is an excellent venue to introduce previously unengaged students to the STEM fields and highlight the importance of scholastic success for athletic achievement. Contrary to previously described sports-science based STEM programs, this program was naturally situated within existing sports programs, allowing STEM to be used by both coaches and players as a tool for athletic improvement. This allows students who would not usually seek out STEM enrichment opportunities to be engaged in STEM within an activity where they are intrinsically motivated to improve. As persistence within the STEM fields has been associated with self-identification as scientists and motivation for learning, the connection between scientific inquiry and basketball training can serve as an arena for this process to take place for typically unengaged students.

We observed a statistically significant increase in both students self efficacy and interest regarding basketball training and STEM attitudes through the use of basketball analytics as a training tool. These preliminary results provide evidence that the application of analytics within youth sports has the potential to serve as both a training tool and as an important, tangible introduction to STEM within a popular youth activity. By expanding the sense of ownership and expertise that a youth feels for basketball into STEM topics through this process, traditionally marginalized students can be brought into the STEM disciplines. As many NBA franchises have already developed extensive youth outreach programs and analytics departments, the use of basic analytical techniques within youth outreach programs is an achievable and impactful method for community enrichment. Similarly, this work provides a framework for the creation of collaborations between university researchers and local youth sports programs for the development of further STEM informed training programs as a method to augment existing university STEM outreach programs.

Acknowledgements

We would like to thank 4th Family Inc. for assistance in developing, coordinating, and deploying the shooting clinic in the community. We would also like to thank Dr. Syb Jennings for assistance with program evaluation. Finally, we would like to acknowledge NSF grant funding: DGE-0947980.

References

- Allen, I. E., & Seaman, C. a. (2007). Likert scales and data analyses. *Quality Progress*, 40(7), 64–65. <http://doi.org/10.1111/j.1365-2929.2004.02012.x>
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37(2), 122–147. <http://doi.org/10.1037/0003-066X.37.2.122>
- Bell, P., Lewenstein, B., Shouse, A. W., & Feder, M. a. (2009). *Science Learning in Designed Settings. Learning Science in Informal Environments: People, Places, and Pursuits*. <http://doi.org/10.1080/00958964.2011.623734>
- Drazan, J. F., Cooke, L., & Eglash, R. (2016). Harmonious Integration : Tuning STEM Education with Generative Justice. In *Integrated STEM Education Conference (ISEC), 2015 IEEE* (pp. 1–8).
- Eglash, R., & Garvey, C. (2014). Basins of Attraction for Generative Justice. *Chaos Theory in Politics*, 1–15.
- Franks, A., Miller, A., Bornn, L., & Goldsberry, K. (2015). Counterpoints : Advanced Defensive Metrics for NBA Basketball. *MIT Sloan Sports Analytics Conference*, 1–8.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference*.
- Germann, P. (1988). Development of the attitude toward science in school assessment and its use to investigate the relationship between science achievement and attitude toward science in school. *Journal of Research in Science Teaching*, 25(8), 689–703.
- Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach’s alpha reliability coefficient for Likert-type scales. *Midwest Research to Practice Conference in Adult, Continuing, and Community Education*, (1992), 82–88. <http://doi.org/10.1109/PROC.1975.9792>
- Goldsberry, K. (2012). Courtvision: New visual and spatial analytics for the NBA. *Proc. 6th Annual MIT Sloan Sports Analytics Conference*, 1–7. Retrieved from http://www.sloansportsconference.com/wp-content/uploads/2012/02/Goldsberry_Sloan_Submission.pdf
- Gottfried, M., & Williams, D. N. (2013). STEM Club Participation and STEM Schooling Outcomes. *Education Policy Analysis Archives*, 21(79), 1–27.
- Karp, T., & Schneider, A. (2011). Evaluation of a K-8 LEGO robotics program. *Proceedings - Frontiers in Education Conference, FIE*, (September). <http://doi.org/10.1109/FIE.2011.6142977>
- Landivar, L. C. (2013). *Disparities in STEM Employment by Sex , Race , and Hispanic Origin. American Community Survey Reports*.
- Lee, X. (2015). *The lack of motivation to pursue postsecondary education among Hmong students: A grounded theory study*.
- Maltese, A. V., Melki, C. S., & Wiebke, H. L. (2014). The Nature of Experiences Responsible for the Generation and Maintenance of Interest in STEM. *Science Education*, 98(6), 937–962. <http://doi.org/10.1002/sce.21132>

- Melchior, A., Cohen, F., Cutter, T., & Leavitt, T. (2005). *More than robots: An evaluation of the FIRST Robotics Competition participant and institutional impacts.*
- NCAA Research. (2016). *Estimated Probability of Competing in College Athletics.*
- President, E. O. of the. (2010). *President's Committee of Advisors on Science and Technology: Panel on Educational Technology, Prepare and Inspire: K-12 Education in Science, Technology, Engineering, and Math (STEM) for America's Future. President's Council of Advisors on Science and Technology.*
- Rothstein, R. (2013). Why Children from Lower Socioeconomic Classes, on Average, Have Lower Academic Achievement Than Middle-Class Children. In *Closing the Opportunity Gap* (pp. 61–74). Oxford University Press. <http://doi.org/10.1093/acprof:oso/9780199982981.003.0005>