Edited by:
Patrick Lambrix
Mikael Vernblom
Niklas Carlsson
Tim Brecht

LINHAC

Linköping Hockey Analytics Conference
LINHAC 2024

LINKÖPING UNIVERSITY

LINKÖPING

Organizing Committee

Patrick Lambrix, Linköping University, Sweden (chair)
Mikael Vernblom, Linköping Hockey Club, Sweden (co-chair)
Niklas Carlsson, Linköping University, Sweden (co-chair)
Tim Brecht, University of Waterloo, Canada (co-chair)

Program Committee for the Research Track

Tim Brecht, University of Waterloo, Canada (chair)
Niklas Carlsson, Linköping University, Sweden (chair)
Martin Arlitt, University of Calgary, Canada
David A Clausi, University of Waterloo, Canada
Jack Davis, University of Waterloo, Canada
Ulf Johansson, Jönköping University, Sweden
Patrick Lambrix, Linköping University, Sweden
Joshua Pohlkamp-Hartt, Boston Bruins, USA
David Thomas Radke, Chicago Blackhawks, USA
Michael Schuckers, Statistical Sports Consulting and St. Lawrence University, USA
Mikael Svarén, Dalarna University, Sweden
Andrew C Thomas, SportsMEDIA Technology, Canada
Sam Ventura, Buffalo Sabres and Carnegie Mellon University, USA
Erik Wilderoth, Färjestad BK, Sweden

In cooperation with

E.H.C. Alliance of European Hockey Clubs

The City of Linköping

In cooperation with

Sportlogiq

Sponsor

The Swedish Research Council for Sport Science

Gold industry collaborator

KINEXON Sports / Dartfish

Silver industry collaborator

Hudl

Silver industry collaborator

Stretch On Sense AB

Bronze industry collaborators

49ing

Spiideo

Sportcontract

Preface

LINHAC 2024 took place June 3-5, 2024, and was organized by Linköping University and Linköping Hockey Club. LINHAC brought together professionals and academics with an interest in hockey analytics. It featured the latest research in hockey analytics in academia and companies, discussions with analysts and coaches, industry sessions with the latest hockey analytics products, and an analytics competition for students.

In addition to the research track, the program included talks by Robin Schuermann (d-fine) on operationalizing analytics for clubs, by Anders Norén from Capacio with Staffan Olsson on cognitive assessment and profiling for increased understanding of game intelligence and performance, by Fredrik Sjöö from Onsite sport on using neuroscience to design a sport's fan-engagement platform, by Lars Skytte on visualizing play-by-play data, by Daniel Weinberger from Sportcontract on constructing a Wins Above Replacement model, by Erik Lignell from Frölunda HC about his experience of a decade as analyst in Swedish elite hockey, and by Craig Buntin from Sportlogiq about the history and goals of Sportlogiq. There were also mini-talks about AI and its possible use in ice hockey analytics. Patrick Lambrix gave a short overview of the field. Albin N Maelum from Stretch on Sense presented an industry perspective. Huanyi Li from Linköping University showed the use of knowledge representation for integration and search of data, and Marco Kuhlmann from Linköping University discussed the use of large language models for user interfaces to databases.

Further, there were panel discussions moderated by Mike Helber and Tim Brecht. A first panel was made up of analysts from different European teams (Miika Arponen from Ässät Pori, Adam Albelin from the Swedish Women's National Team, Zach Ellentahl from Rögle BK, Jan Morkes from Bílí Tygři Liberec and the Czech Men's National Team, and Simo Teperi from Rauman Lukko). The second panel discussed the state of the art and future of hockey analytics from the industry perspective (Thorsten Apel from Sportcontract, Lance Du'Lac from Hudl, Michael Elmer from KINEXON Sports, Andreas Hänni from 49ing, Albin N Maelum from Stretch on Sense, Jean-Sébastien Mérieux from Dartfish, and Morgan Zeba from Spiideo). The third panel discussed hockey analytics in the public sphere and how to get into the industry and teams. The participants were Petter Carnbro from Leksands IF, Lars Skytte (`https://hockey-statistics.com/`), Daniel Weinberger from Sportcontract, and Erik Wilderoth from Färjestad BK. Finally, the entertainment industry panel discussed their use of analytics (Almen Bibic from TV4, Meghan Chayka from Statlethes, Andreas Hänni from 49ing, and Henric Larsson from TV4/Hockeylabbet).

There were also three interviews with analysts from NHL teams about the state-of-the-art use of hockey analytics. The first interview was with Josh Pohlkamp-Hartt and Campbell Weaver from the Boston Bruins, the second with Katerina Wu from the Pittsburgh Penguins, and the third with Dave Radke from the Chicago Blackhawks.

Our industry collaborators presented their products: KINEXON Sports / Dartfish, Hudl, Stretch on Sense, 49ing, Spiideo and Sportcontract.

Finally, there was a student competition where the task was to provide insights based on sequences of events in a hockey game. Data was provided by the SHL and Sportlogiq.

LINHAC is the only conference of its kind in Europe, and as far as we know, it is the only hockey analytics conference that covers all aspects related to hockey analytics. This book includes the research track proceedings as well as contributions from industry, the student competition papers, and insights from contributors to LINHAC regarding their experience with hockey analytics and thoughts about its future. The research track proceedings are also published at Linköping Electronic Conference Proceedings as ECP 209, `https://doi.org/10.3384/ecp209`.

We thank our moderator Mike Helber, our conference service TM Event, and the members of our local organization committee Mina Abd Nikooie Pour, Huanyu Li, Ying Li, Gurjot Sing, Chunyan Wang, Jenny Rydén, Lene Rosell, Anders Cronstierna, and Daniel Jemander, for their excellent support.

Last, but not least, we thank our collaborators the Alliance of European Hockey Clubs, the City of Linköping, Sportlogiq, our sponsor the Swedish Research Council for Sport Science, and our gold (KINEXON Sports / Dartfish), silver (Hudl, Stretch on Sense) and bronze (49ing, Spiideo, Sportcontract) industry collaborators.

September 2024

Patrick Lambrix (chair),
Mikael Vernblom (co-chair),
Niklas Carlsson (co-chair),
Tim Brecht (co-chair)

Contents

Invited talk

Cognitive Assessment and Profiling for increased understanding of Individual and Team Game Intelligence and Performance in Ice hockey

Anders Norén

Capacio, Sweden

Abstract. Game intelligence, the ability to be in the right place at the right time and make optimal decisions, is crucial for athletic performance. This whitepaper explores how neurocognitive testing and profiling can deepen our understanding of game intelligence, which includes elements such as situational awareness, decision-making, problem-solving, and flexibility.

The whitepaper targets sports professionals aiming to enhance their understanding of game intelligence through neurocognitive assessments. The assessments mentioned in the paper provide insights into athletes' cognitive strengths and weaknesses, aiding in talent identification, personalized coaching, strategic team composition, tactical adaptations, and injury prevention. Executive functions are crucial in both open sports (e.g., soccer, basketball) and closed sports (e.g., archery, golf). For example, in ice hockey, players must continuously adapt to dynamic environments, requiring quick decision-making, strategic thinking, and creativity.

Integrating neurocognitive assessments into sports practices has the potential to enhance the understanding of game intelligence, reduce subjectivity and bias, and improve individual and team performance, As well as ensure the wellbeing of athletes through tailored mental health support and coping strategies. Testing and profiling of individuals and teams can practically help enhance understanding of Game Intelligence. The process involves assessment, awareness, individual acceptance, strategic development, integration into coaching, and continuous follow-up to monitor progress and aid adjustments.

Game Intelligence in Sports

Game intelligence, the ability to be in the right place at the right time and make the right decisions, is a critical aspect of athletic performance. This whitepaper explores how neurocognitive testing and profiling can deepen our understanding of game intelligence, encompassing elements such as situational awareness, decision-making, problem solving and adaptability.

Understanding Game Intelligence

Game intelligence refers to the cognitive skills and processes that enable athletes to anticipate play developments, make strategic decisions, and execute actions effectively under varying conditions. Recent scientific studies indicate that distinct aspects of game intelligence can be predicted by examining specific executive functions (EFs) such as attention, working memory, cognitive flexibility, creativity, and impulse control.

Defining Game Intelligence and Executive Functioning

- **Game Intelligence**: The ability to understand and anticipate game situations, make strategic decisions, and execute appropriate actions. Or to be at the right place at the right time and do the right thing.
- **Executive Functioning** : The cognitive processes enabling individuals to focus on the right things and perform successfully in constantly changing environments. This ability varies between individual, and depend on a combined set of cognitive functions that interact.

These definitions highlight the executive functioning are central in forming game intelligence. Measuring and understanding executive functions can hence help understand distinct aspects of game intelligence, create insights, and aid us in how to coach and apply efficient strategies that support an athlete's performance and development.

Target Audience

This whitepaper is intended for sports professionals seeking to enhance their understanding of game intelligence, including scouts, sports directors, coaches, general managers, mental coaches, physiotherapists, and athletes. By integrating neurocognitive assessments into their talent identification, development and coaching processes, these professionals can gain a more comprehensive understanding of an athlete's potential and performance capabilities and integrate this knowledge to better support the development of individual athletes. As well of the optimal function of a team.

The Role of the Brain in Game Intelligence

The Brain and in particular Executive Functions in the frontal lobe are crucial for controlling and regulating information processing, thoughts, emotions, and performance. They include among others:

- **Attention**: Focusing on relevant information while ignoring distractions.
- **Short Term Memory**: Retaining and utilizing information over brief periods.
- **Working Memory**: Holding and manipulating information over short periods.
- **Impulse Control**: Suppressing inappropriate actions or responses.
- **Cognitive Flexibility**: Adjusting to changing situations and demands.
- **Creativity**: Generating novel ideas and solutions.

- **Conceptualization**: Forming and manipulating concepts.
- **Strategic Thinking**: Planning and executing strategies.

There are significant differences in cognitive functioning between individuals (and athletes). By assessing these differences, we can create cognitive profiles, where strengths and weaknesses of the individual can help us understand aspects of game intelligence and how to work with individual strengths and potential while finding efficient strategies to compensate for weaknesses. Based on the athletes' different cognitive profiles, it is possible to apply strategies that improve both individual and overall performance of teams. It is also possible to improve well-being among athletes by understanding and basing coaching and the athletes' roles of individuals cognitive profiles.

Measuring Cognition and Executive Functioning

Cognitive neuroscience has existed as a science for more than 80 years with the purpose of better understanding individuals based on their inherent abilities and help optimize function. The last couple of decades research has spread from focusing on function deficit to also study talent and top performance, particularly in sports.

Neurocognitive tests have been developed to measure these abilities and can be used to provide a detailed profile of an athlete's cognitive strengths and weaknesses. Cognitive capacities in executive functions are intricately linked to an individual's executive performance, which refers to their ability to function effectively in dynamic and demanding environments. Research constantly indicates that the validity of these measurements is high, ranging from 70-90% when it comes to measuring the capacity of these functions, on both general and athletic populations.

Application in Sports - Talent Identification and Development

By using neurocognitive profiling, sports organizations can better understand an athlete's cognitive capacities, which are crucial for game intelligence. These insights can facilitate better decisions related to scouting, team composition, and both personalized and situational tactics and coaching strategies for individuals and teams. An increased self-awareness and understanding of others can also improve team dynamics and collaboration.

Open and Closed Sports

Open sports are activities where the environment is unpredictable, and athletes must respond to changing conditions and the actions of opponents. Characteristics:

- **Unpredictable Environment**: Conditions change constantly, requiring quick adjustments.

- **Reactive**: Athletes must react to opponents, teammates, and environmental factors.
- **Examples**: Soccer, basketball, tennis, rugby, hockey.

Closed sports are activities performed in a stable, predictable environment where the athlete initiates the action. Characteristics:

- **Predictable Environment**: Conditions remain stable, allowing for consistent execution.
- **Self-Paced**: Athletes control the timing and pace of their actions.
- **Examples**: Archery, golf, bowling, gymnastics routines, diving, target shooting.

Key Differences between Open and Closed Sports:

- **Environment**: Open sports have dynamic and unpredictable environments, while closed sports have stable and predictable ones.
- **Action**: Open sports require reactive skills, while closed sports rely on preplanned, self-paced actions.

General examples of Executive Functions in Open Sports

- **Attention in High-Pressure Situations**: Studies show that high-level attention is critical in sports where multiple events occur simultaneously, such as basketball soccer, hockey etc. Enhanced situational awareness allows athletes to perceive and respond to dynamic play developments effectively.
- **Working Memory in Complex Play**: In sports like soccer, handball, football, basket, baseball etc, athletes need to hold information, such as instructions, online, and act on rapidly on them. Working memory facilitates the retention and application of these instructions during fast-paced games, for example adaption and integration of the instructions with the ongoing game.
- **Cognitive Flexibility and Creativity in Soccer**: Midfielders and forwards, especially in soccer and ice hockey, require elevated levels of creativity and cognitive flexibility to switch between offensive and defensive roles seamlessly and generate novel multi-step solutions to emerging challenges.
- **Speed vs. Accuracy in Decision-Making**: The balance between speed and accuracy varies across sports. For instance, in football and ice hockey, quick decision-making often takes precedence over precision due to the fast-paced nature of the games. Hence the focus on ball possession and retaking the ball. In contrast, handball emphasizes accuracy, as the sport involves more structured play, and the consequences of mistakes are magnified. Hence the focus on not losing the ball.
- **Impulse Control in Timing Actions**: An athlete's ability to regulate and time their actions precisely under pressure is vital. For example, a soccer player may need to control their impulse to shoot immediately and instead

wait for the optimal moment, ensuring a higher success rate. Especially defenders or goal keepers benefit from a strong impulse control not to be lured by an opponent, a process that requires behavioural inhibition. The timing is crucial and missing the action may result in a goal and a lost match.
- **Strategic Thinking in Pre-Planned Plays**: Sports like American football, basketball, and handball often involve complex, pre-planned plays. Athletes must learn and execute these plays accurately, which requires strong strategic thinking and memory skills combined.
- **Variability in Performance**: An athlete's executive functioning can fluctuate based on conditions such as fatigue and stress. For example, a player's impulse control might be high in calm situations but diminish under pressure, affecting their timing and emotional regulation. Understanding individual sensitivity to changes within and around may allow athletes to optimize in a decisive way to be at their best when it matters.

Importance of Executive Functions in Ice Hockey

Ice hockey is a fast-paced, high-stakes sport where players must continuously adapt to the dynamic environment. The role of executive functions in ice hockey is critical due to the need for quick decision-making, strategic thinking, and effective communication.

- **Focus and Attention**: During a power play, a player must maintain focus on the puck while being aware of the positions of teammates and opponents. This requires sustained attention and the ability to filter out irrelevant stimuli.
- **Cognitive Flexibility**: A defenseman might initially plan to pass the puck to a teammate on the left, but if an opponent intercepts that path, they need to quickly switch strategies and find an alternative passing route or decide to clear the puck.
- **Creativity**: Players must employ creative thinking to devise unexpected plays and manoeuvres, such as innovative passing sequences, deceptive shots, and creative dekes to outmanoeuvre opponents and create scoring opportunities.
- **Inhibition Control**: Controlling impulses is vital during face-offs and when checking opponents. A player must avoid unnecessary penalties by restraining from actions like high-sticking or tripping, which could harm the team.
- **Working Memory**: Players need to remember and execute complex play formations, such as a breakout strategy from the defensive zone. They must also recall the tendencies and strategies of the opposing team from pre-game analysis.
- **Planning and Strategic Thinking**: For a forward breaking into the offensive zone, planning involves deciding whether to pass, shoot, or deke based on the positions of the defenders and the goalie. They must anticipate the possible reactions of their opponents and teammates.

- **Team Dynamics and Communication**: Effective communication and understanding of team dynamics are crucial. Players must quickly interpret and respond to verbal and non-verbal cues from teammates to execute plays successfully, such as a quick pass during a fast break or coordinating a defensive strategy to counter an opponent's attack.

Examples of Executive Functions in Closed Sports

Closed sports, also known as self-paced sports, are activities where the athlete initiates the action and performs in a stable and predictable environment. Understanding and controlling executive functions in these sports is crucial due to the need for precision, concentration, and strategic planning. Here are some examples of closed sports along with the challenges where executive functions are essential:

Archery

- **Focus and Attention**: Archers need to maintain intense concentration for extended periods to aim and release the arrow accurately.
- **Inhibition Control**: The ability to suppress distracting thoughts and external noises is crucial to maintain a steady hand and precise aim.
- **Cognitive Flexibility**: Adjusting to varying wind conditions or slight changes in lighting without losing focus.

Golf

- **Planning and Strategy**: Golfers must plan their shots carefully, considering factors like wind, terrain, and distance.
- **Working Memory**: Remembering previous shots, course layout, and adjusting technique accordingly.
- **Inhibition Control**: Managing frustration and maintaining composure after a poor shot to avoid affecting subsequent shots.

Bowling

- **Attention to Detail**: Bowlers need to focus on their approach, timing, and release to ensure accuracy and consistency.
- **Cognitive Flexibility**: Adjusting technique based on lane conditions and performance of previous frames.
- **Inhibition Control**: Controlling emotions and staying calm under pressure, especially in competitive settings.

Gymnastics (certain routines)

- **Inhibition Control**: Suppressing nervousness and distractions to perform complex sequences accurately.

- **Focus and Attention**: Maintaining attention on precise execution of movements while ignoring external stimuli.

Target Shooting (e.g., rifle or pistol shooting)

- **Focus and Attention**: Shooters need to maintain elevated levels of concentration to aim and fire accurately.
- **Inhibition Control**: Suppressing physical and mental distractions, such as muscle tremors or anxiety.
- **Cognitive Flexibility**: Making minute adjustments to aim based on changing environmental conditions.

Diving

- **Planning and Strategy**: Divers need to plan their dives meticulously, considering the sequence of movements and entry into the water.
- **Working Memory**: Remembering and executing complex routines.
- **Inhibition Control**: Managing stress and maintaining composure before and during the dive to ensure precise execution.

Billiards/Pool

- **Strategic Thinking**: Planning shots several moves ahead to control the table.
- **Attention to Detail**: Precise control over cue ball and object balls requires intense focus.
- **Inhibition Control**: Maintaining calmness and control, especially in high-pressure situations.

Cycling

- **Focus and Attention**: Cyclists need to maintain focus on their pace, breathing, and positioning on the bike for extended periods. Any lapse in concentration can lead to decreased performance or accidents.
- **Cognitive Flexibility and quick decision making**: During a race, a cyclist may need to adjust their strategy based on changing weather conditions or the behaviour of other competitors. If a rider breaks away from the pack, the cyclist must quickly decide whether to follow or stick to their planned pace.
- **Inhibition Control**: Cyclists must manage their energy output and resist the urge to push too hard too early in the race, which could lead to burnout. They need to stick to their planned strategy and pacing.
- **Working Memory**: Remembering the course layout, including the locations of steep climbs, sharp turns, and aid stations, is crucial for effective race management. This allows for strategic energy conservation and optimal performance.
- **Planning and Strategic Thinking**: Effective race planning involves setting a pace strategy that considers the cyclist's strengths and the course profile. For example, knowing when to conserve energy on flats and when to push hard on climbs can make a significant difference in overall performance.

Importance of Executive Functions

In all sports, the ability to control executive functions directly impacts performance. To control we need to understand and therefore we need to measure. When measuring it is also important to examine and related capacity to awareness and existing coping strategies that might already be in place.

Based on enhanced awareness, effective adaption of training and coaching can help athletes enhance their performance by improving their mental control, decision-making, and overall strategic approach to their sport. As always in performance-intense environments slight changes may have a significant impact on the result.

Practical Implications

Integrating neurocognitive assessments into sports practices provides a robust basis for:

- **Enhanced Understanding of Game Intelligence**: Improving overall game intelligence by understanding how cognitive functions influence decision-making and situational awareness.
- **Reduce Subjectivity and Bias**: Providing objective data to minimize biases in talent identification, training, and team selection decisions.
- **Enhanced Understanding of Individual Functioning**: Gaining deeper insights into how each athlete's cognitive abilities impact their performance and tailoring support accordingly.
- **Enhanced Talent Identification**: Identifying athletes with high potential based on cognitive profiles.
- **Personalized Coaching**: Developing tailored training programs to, based on cognitive profile, strengths and weaknesses, support development of optimizing and compensating strategies as well as adapt physical, technical, tactical, and mental training.
- **Strategic Team Composition**: Forming teams with complementary cognitive strengths.
- **Tactical and Strategic Adaptations**: Adapting tactics and strategies to leverage cognitive strengths and address weaknesses, both at individual and team levels.
- **Matching Players**: Pairing or grouping players based on compatible cognitive profiles to enhance on-field synergy and effectiveness.
- **Co-Play and Team Dynamics**: Fostering better teamwork and communication by understanding and leveraging cognitive dynamics within the team.
- **Time Efficiency**: Saving time by quickly understanding athletes' cognitive profiles, allowing for faster and more effective training adaptations and decision-making processes.
- **Stress resilience and Well-being**: Enhancing athlete well-being by recognizing cognitive stressors, providing appropriate mental health support, and coping strategies.

- **Injury Prevention and Assessment**: Utilizing cognitive assessments to prevent injuries, including concussions, by identifying risk factors and monitoring cognitive health, and assessing cognitive impact post-injury to guide rehabilitation and return-to-play decisions.

Implementing Understanding of Executive Functioning

- **Assessment and Profiling**: In just 45 minutes, an athlete's cognitive profile can be assessed, providing a robust evidence base for their development. This quick yet comprehensive assessment ensures timely and accurate understanding of cognitive strengths and weaknesses. And can save weeks, months or even years of struggling to coaching and development. Thus supporting a quick transition from Good to Great.
- **Understanding and Awareness**: Fill the gap by understanding the "how" and "why" behind observed performance and behaviour. Engage all relevant stakeholders to ensure an integrated approach, enhancing cooperation and minimizing misunderstandings.
- **Acceptance and Commitment**: Ensure the individual athlete's acceptance of the test results and foster their commitment to working with the insights gained. This step is crucial for effective application and long-term success. Make sure all stakeholders are involved and have a sufficient level of understanding.
- **Optimizing and Compensating Strategies**: Leverage the cognitive profile to identify effective coping strategies not just cognitively, but also across the tactical, physical, technical, and mental domains. This multi-faceted approach ensures comprehensive development and performance optimization.
- **Integration into Existing Coaching and Training Schemes**: Seamlessly integrate the cognitive insights into daily coaching and training routines. This ensures that all involved actors, including coaches and support staff, are aligned and can work synergistically towards the athlete's development.
- **Follow-Up and Adjustment**: Establish a continuous follow-up mechanism to monitor progress and make necessary adjustments. This ensures sustained understanding, reinforcement of strategies, and ongoing development.

References

1. Vestberg, T., Gustafson, R., Maurex, L., Ingvar, M., Petrovic, P. (2012). "Executive functions predict the success of top-level football players". PLoS ONE. `https://doi.org/10.1371/journal.pone.0034731`
2. Vestberg, T., Reinebo, G., Maurex, L., Ingvar, M., Petrovic, P. (2017). "Core executive functions are associated with success in young elite soccer players". PLoS ONE. `https://doi.org/10.1371/journal.pone.0170845`
3. Vestberg, T., Gustafson, Jafari, R., Almeida, R., M., Maurex, L., Ingvar, M., Petrovic, P. (2020). "Level of play and coach-rated game intelligence are related to performance on design fluency in elite soccer players". Scientific Reports. `https://doi.org/10.1038/s41598-020-66180-w`

4. Blecharz, J., Wrześniewski, K., Siekańska, M., Ambroży, T., Spieszny, M. (2022). "Cognitive Factors in Elite Handball: Do Players' Positions Determine their Cognitive Processes?" Journal of Kinetics. https://doi.org/10.2478/hukin-2022-0063
5. Verburgh, L., Scherder, E. J. A., van Lange, P. A. M., Oosterlaan, J. (2014). "Executive functioning in highly talented soccer players". PLoS ONE. https://doi.org/10.1371/journal.pone.0091254
6. Tamas, B., Géczi, G., Gurisatti, L. (2024). "Executive functions and competitive attitudes in near-elite ice hockey players". Physical Culture and Sport. Studies and Research. https://doi.org/10.2478/pcssr-2024-0015
7. Glavas, D., Pandzic, M., Domijan, D. (2023). "The role of working memory capacity in soccer tactical decision making at different levels of expertise." Cognitive Research Principles and Implications. https://doi.org/10.1186/s41235-023-00473-2
8. Scharfen, H.E., Memmert, D. (2019). "The Relationship Between Cognitive Functions and Sport-Specific Motor Skills in Elite Youth Soccer Players." Frontiers in Psychology. https://doi.org/10.3389/fpsyg.2019.00817
9. Montuori, S., D'Aurizio, G., Foti, F., Liparoti, M., Lardone, A., Pesoli, M., Sorrentino, G., Mandolesi, L., Curcio, G., Sorrentino, P. (2019). "Executive functioning profiles in elite volleyball athletes: Preliminary results by a sport-specific task switching protocol." Human Movement Science. https://doi.org/10.1016/j.humov.2018.11.011
10. Yongtawee, A., Park, J., Kim, Y. Woo, M. (2022). "Athletes have different dominant cognitive functions depending on type of sport." International Journal of Sport and Exercise Psychology. https://doi.org/10.1080/1612197X.2021.1956570
11. Mansi, E., Gunay, E., Guducu, K., Herold, F. (2023). "The Effect of the Playing Positions in Basketball on Measures of Cognitive Performance". Journal of Cognitive Enhancement. https://doi.org/10.1007/s41465-023-00269-8
12. Németh, K., Balogh, L. (2021). "The relationship between decision-making and game intelligence with basketball statistics." Hungarian Journal of Sport Sciences. https://doi.org/10.36439/shjs/2021/1/9535

Research papers

Evaluating Space Creation in the National Hockey League using Puck and Player Tracking Data

Hassaan Inayatali and Timothy Chan

University of Toronto

Abstract. Star ice hockey players are often described as having a magnetic pull, with the ability to draw out opponents and generate dangerous opportunities for their linemates in the space left vacant by defenders. Using spatiotemporal Puck and Player Tracking (PPT) data, we develop a quantitative approach to measure how players create space while in possession of the puck, termed On-Puck Space Generation (OPSG). The benefits of our model's approach include its decomposition into three components: 1) Rink Control, the probability of controlling the puck at a given location; 2) Rink Value, the probability of scoring from a given location; and 3) Transition Probability, the probability that the next on-puck event will occur at a given location. Preliminary results of our metric show that players who achieve high levels of OPSG are more likely to lead their team in goals, assists and points. Our model can be used to analyze which players are in positions of danger, identify instances in which an individual created valuable space for their teammates, and understand which teams are best at generating space.

1 Introduction

While much of the information used to construct ice hockey teams and evaluate players is limited to the contributions of an individual, success in ice hockey requires high degrees of coordination among teammates. A common point of discussion with regards to play-making is space creation, movement which enables fellow teammates to position themselves in areas of high value. The work in this paper aims to address the following research question: How can we quantify the value of player movement with respect to influencing defender actions and creating scoring opportunities? In this work, we develop a model to quantify space creation by players in possession of the puck.

In the past, ice hockey analytics have been limited to event and stint data, which includes actions on the puck, the players involved and which players were on the ice at the time of a given event. Building off of the work by Sam Green [5] in soccer, expected goals (xG) models were developed to better understand the quality of shot opportunities in ice hockey [7, 3]. More comprehensive models such as Regularized Adjusted Plus-Minus (RAPM) models [4] have been built to better understand an individual's offensive and defensive impact. However, these models were unable to include information on teammate and opponent positioning on the ice.

With the introduction of Puck and Player Tracking (PPT) Data during the 2019-2020 NHL Playoffs, we can obtain more context into the game state when an event occurs and answer more complex questions relating to space creation. With other sports like soccer having had access to spatiotemporal data for nearly a decade [6], researchers have developed methodologies which focus specifically on inter-player dynamics. Thus, we have the opportunity to adapt existing models to ice hockey while accounting for the differences in these sports.

From this literature, we develop a model to determine the probability of scoring on the next on-puck event given the state of the game conditioned on an instantaneous pass event. Our model consists of three components:

1. **Rink Control**: The probability of controlling the puck at a given location.
2. **Rink Value**: The probability of scoring from a given location.
3. **Transition Probability**: The probability the next on-puck event will occur at a given location.

We calculate the change in the probability of scoring from the start of the possession to the end and aggregate across all possessions. This yields our final metric, On-Puck Space Generation (OPSG). The results of our model can help in understanding specific instances of space creation, evaluate player movements, and discern which teams are able to generate space consistently against their opponents. The contributions of this work are as follows:

– We develop a novel transition probability model for ice hockey conditioned on the locations and movements of all players on the ice.
– We propose a model to measure space creation by players in possession of the puck. Our model is composed of three sub-models which increase interpretability of the model's predictions.
– We perform an evaluation over 35 NHL games from the 2023-2024 season. We aggregate OPSG for individual players and teams. Our results show that OPSG has the strongest correlations with forward assists, defensemen goals, and team shot attempt differential.

2 Related Work

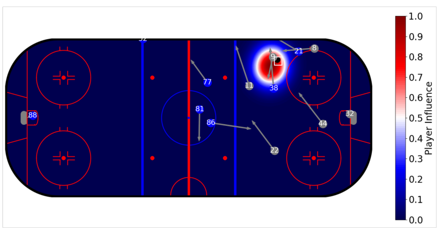
Pitch control refers to the "the probability that a player or team will be able to control the ball if it were at that location" [11]. Pitch control models have been developed in various forms, through the use of Voronoi Diagrams [12], Player Influence Models [1], and Poisson Point Processes [10]. While Pitch Control provides insight into spatial ownership, the value of this space is not considered. Pitch Value models aim to learn the value of space in different areas of a playing surface. These models can apply defensive positioning [1], distance from the net [10], or other models to decompose possession value into various actions [2]. These models are combined to better understand the quality of space controlled by each team [1, 10]. These can be used to create off-ball space creation metrics including Off-Ball Scoring Opportunity (OBSO) [10] and Space Generation Gain (SGG) [1]. In ice hockey, past research with PPT Data have focused on passing lanes [8] and passing value [9].

In this paper, we adapt Pitch Control and Pitch Value as well as expected pass speed from the aforementioned passing lane literature [8] to ice hockey to construct each of our Rink Control, Rink Value, and Transition Probability models. Our final result is a novel metric, entitled On-Puck Space Generation (OPSG), which examines a player's creation of space while in possession of the puck. We analyze how OPSG correlates with a player's cumulative production in terms of goals, assists, and points. Furthermore, we delve into a team's game-level space creation and how it relates to their performance in terms of shot attempts, shots on goal (SOG) and goals scored. Unlike previous work that focuses explicitly on puck transitions between players, our model provides insight into how players generate value when in possession of the puck. To the best of our knowledge, this is the first such model in the domain of ice hockey.

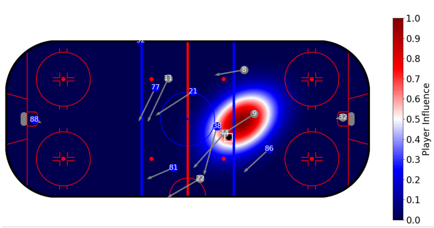
3 Methodology

3.1 Rink Control

We develop our Rink Control model using a bivariate normal distribution, in a similar fashion to Bornn and Fernandez [1]. To account for the speed of ice hockey, we increase the denominator in calculating a player's normalized speed ratio to 1500 ft/s. Additionally, we set the range of influence to be a minimum of 12 ft, which increases with distance from the puck up to a maximum of 30 ft. Distance from the puck affects the range of influence, aligning with the premise of Bornn and Fernandez, that "if the ball moves toward the player he would have more time to reach the ball within a larger space" [1]. Figure 1 shows two examples of player influence for a single player, both with the puck (Figure 1a) and without the puck (Figure 1b), ranging from 0 to 1. We focus on #9 in grey, where darker shades of red indicate higher levels of influence. For this and all future figures, the white square highlights the puck's location.



(a) Player Influence with the Puck



(b) Player Influence without the Puck

Fig. 1. Visualizations for Continuous Player Influence

The influence of each player is aggregated and the home team's influence is subtracted from the visitor's at each location. The logistic function is applied to obtain a measure of Rink Control for each team, in the range of 0 to 1. A sample of this can be seen in Figure 2. Darker shades of red represent higher levels of influence for the team in gray, whereas darker shades of blue represent higher levels of influence for the blue team.

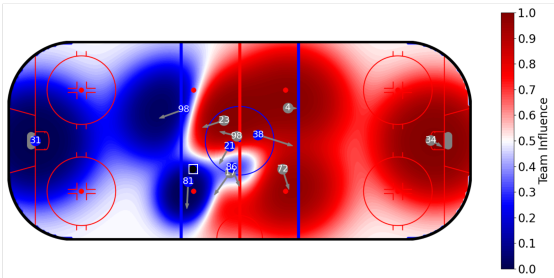


Fig. 2. Team Influence Model

3.2 Rink Value

To determine the value of a region on the ice, we develop an expected goals (xG) model to predict the probability a goal will be scored from a given location. This model is a logistic regression which predicts the probability of scoring based on the distance and angle of a shot. Our model is trained using NHL play-by-play (PBP) data from the 2015-2016 season, and achieves a cross-validated AUC of 0.731. This model can be seen in Figure 3. Darker shades of red represent a higher probability of scoring from the perspective of the team in grey, from 0 to 0.4.

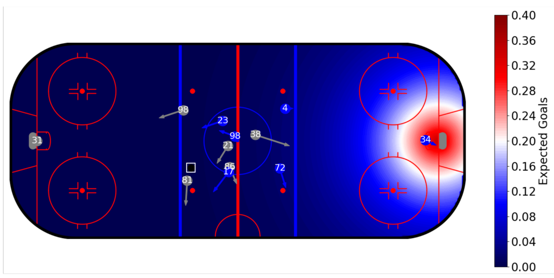


Fig. 3. Expected Goals (xG) Model

3.3 Transition Probability Model

With our Rink Control and Rink Value models, we determine the probability the next on-puck event occurs at a given location. using the following procedure:

1. Obtain a dataset of intended passes (successful and failed).
2. Model the probability a given pass will be successful.
3. Calculate the probability a pass will be successful to any location.
4. Normalize Pass Probability such that the sum over the rink is equal to 1

Thus, we assume Transition Probability is proportional to the probability a pass will be successful to a given location.

Possession Model Given that we do not have access to event data/passes, it is assumed that passes are transfers of possession between teammates. To obtain passing instances, we develop a possession model that produces a binary variable indicating the team in possession of the puck. The rule-based model is as follows:

1. The player is closest to the puck AND
2. The player is within six feet of the puck (one stick length) AND
3. The puck is traveling slower than 35 ft/s (max skater speed) AND
4. (a) The player was the previous player in possession of the puck OR
   (b) The player has an additional six feet between themselves, the puck, and the nearest defender

Pass Regression Using successful passes, we apply the linear regression methodology employed by Radke et al. [8] to predict the time until a pass will arrive at a receiver given their the distance from the passer. Transfers of possession were limited to less than two seconds to ensure only intended passes were analyzed as opposed to dump-ins or puck recoveries. This model was developed using 697 successful passes taken over the course of an NHL game, shown in Figure 4.

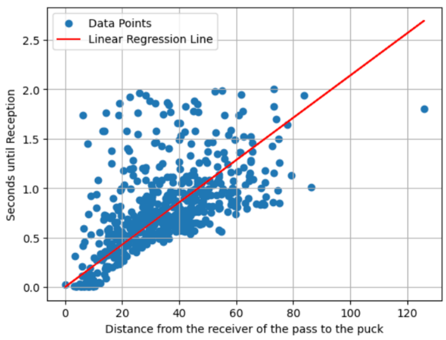


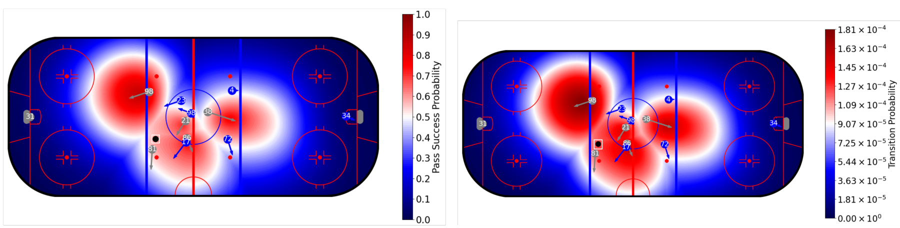
Fig. 4. Passing Linear Regression Model

The slope of this regression is 0.021 s/ft, meaning that for every additional foot a player is away from the puck carrier, it will take an additional 0.021 s for the puck to reach this player.

Pass Prediction To predict passes, we take the velocity vector of the puck at the moment the player no longer has possession, decided by our rule-based model presented in Section 3.3. For each teammate, the puck's velocity is projected according to how far they are from the puck carrier, using the linear regression in Figure 4. Each teammate's velocity is projected forward to estimate where they will be at the puck's arrival [8]. The distance between the puck's projection and teammate's projection is calculated. The player with the smallest distance is predicted to be the receiver of the pass. We exclude passes which hit the boards (dump-ins, bank passes, rims) as well as those which would be further from their intended receiver than 10 feet, to ensure we only analyze passes in which there was the direct intention to be received by a given player. A pass is successful if the following player in possession of the puck is the passer's intended target.

We fit a logistic regression model to predict pass success probability using five variables: defensive influence at the origin, midpoint and destination of the pass, projected distance between reception and receiver, and pass distance. The model was trained on 7000 passes and achieved a cross-validated AUC of 0.751.

Transition Probability We assume that a pass to a given location is intended for the player with the highest probability of receiving it. Because Transition Probability is proportional to the probability a pass would be successful to a given location, we normalize team pass probabilities across the rink surface to sum to 1. Team Pass Probability and Transition Probability can be seen in Figure 5. In Figure 5a, the right-defensemen of the grey team is in possession of the puck and the highest probability of a pass being successful is to his defensive partner. However, there are also passing lanes to each forward, and this can be seen by the darker shades of red in the direction each is travelling. Figure 5b takes the results from Figure 5a and normalizes across the rink surface by dividing the probability at each location by the sum of all probabilities.



(a) Team Pass Probability Model (b) Transition Probability Model

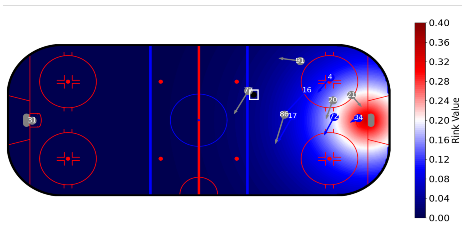
Fig. 5. Visualizations for Continuous Player Influence

3.4 Combined Model

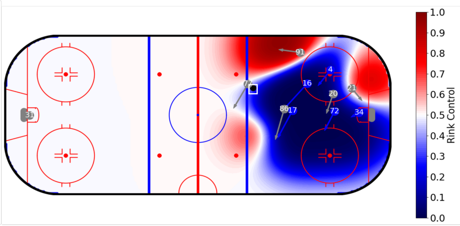
Using the framework presented by Spearman [10], we predict the probability of scoring on the next on-puck event for the attacking team conditioned on the puck possessor passing the puck. Let G_r denote the probability of scoring from location r, C_r represent the probability of controlling the puck at location r, and T_r signify the probability of passing the puck to location r. D and M represent the state of the game and a boolean representing an instantaneous passing event, respectively. Equation 1 shows how these models are combined to calculate the probability of scoring on the next on-puck event for the attacking team given the state of the game conditioned on an instantaneous pass event, $P(G|D, M)$.

$$P(G|D,M) = \sum_{r \in RXR} P(G_r|C_r, T_r, D, M) P(C_r|T_r, D, M) P(T_r|D, M) \tag{1}$$

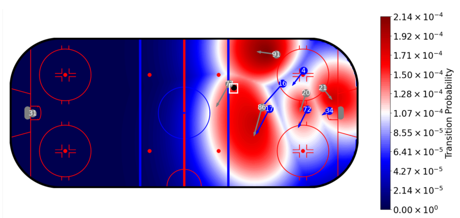
Figure 6 illustrates each component of the model, with Figure 6a showing Rink Value, Figure 6b displaying Rink Control, Figure 6c demonstrates Transition Probability and Figure 6d presents the combined model, with darker shades of red representing higher values. The sum of the combined model across the rink surface is $P(G|D, M)$. In Figure 6d, the probability of scoring is concentrated on the top right of the rink surface, driven by there being a player who is in a position of high value (Figure 6a), able to control the puck if it were to reach them (Figure 6b) and receive a pass at that location (Figure 6c).



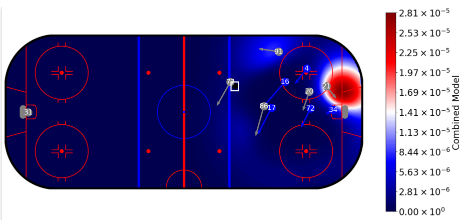
(a) Rink Value Model



(b) Rink Control Model



(c) Transition Probability Model



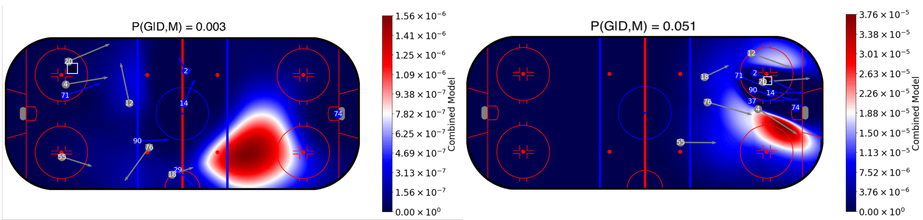
(d) Combined Model

Fig. 6. Visualizations for Continuous Player Influence

4 Results

4.1 Tactical Analysis

Using our model presented in Section 3, we measure how players create valuable space for their teammates. Figure 7 represents one of these situations, when the puck carrier starts with the puck in their own zone and proceeds to carry the puck from end-to-end and creates a direct passing lane to their teammate on a 2-on-1. At the beginning of the possession (Figure 7a), the probability of scoring on the next on-puck event is concentrated for #18 (bottom), which would depend on his being able to beat his defender to receive the pass from #20 (top left). By the end of the possession (Figure 7b), the teammate closest to the puck carrier (#4) on the 2-on-1 is occupying a high value area with a passing lane to receive the puck. This process can support coaches in identifying moments where an individual creates valuable space for their teammates in opposition scouting as well as better understanding how their players create space for one another.

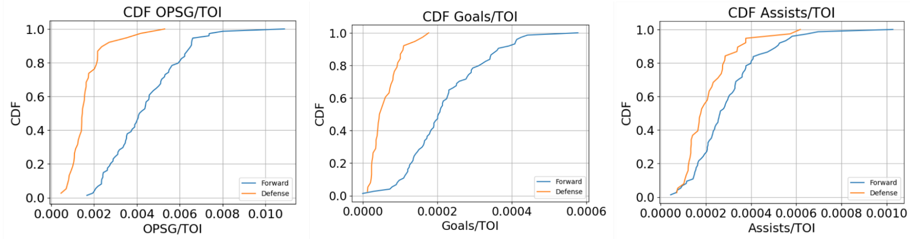


(a) Start of Possession (b) End of Possession

Fig. 7. Visualizations for Continuous Player Influence

4.2 Player Evaluation

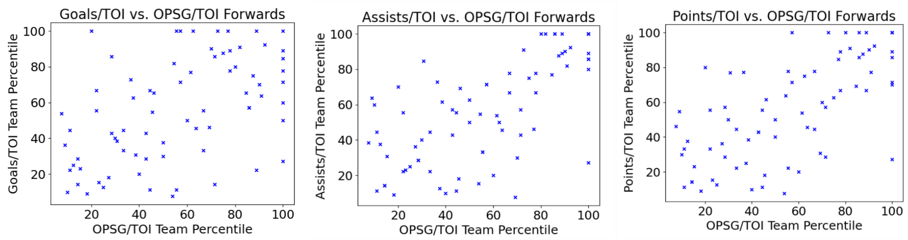
We aggregate the change in probability of scoring on the next on-puck event over a player's possessions to measure their On-Puck Space Generation (OPSG). We complete this process across 35 NHL games from the 2023-2024 season. Matchups were selected to maximize games played (GP) by a subset of teams to gather a representative sample for each player given compute constraints. Our evaluation examines players with five or more GP. This includes 74 forwards and 38 defensemen, with their mean time on ice (TOI) being 16.2 and 20.7 minutes per game, respectively. We normalize metrics using TOI to ensure usage does not affect player comparisons. In Figure 8 we plot CDF's for OPSG/TOI, Goals/TOI and Assists/TOI. Forwards generate more OPSG/TOI than defensemen (Figure 8a), with the 80^{th} percentile forward generating three times more than the 80^{th} percentile defensemen. This is comparable for Goals/TOI (Figure 8b); however the difference for Assists/TOI is considerably smaller (Figure 8c).



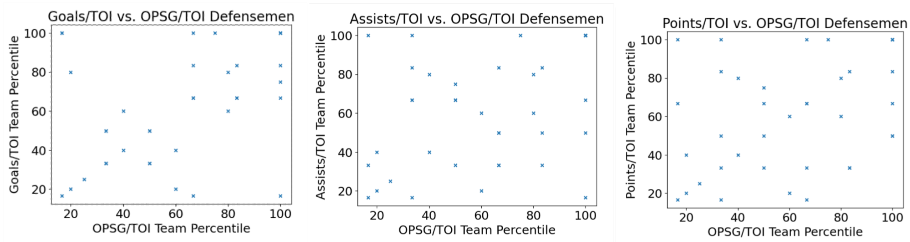
(a) OPSG/TOI (b) Goals/TOI (c) Assists/TOI

Fig. 8. CDF Plots by Position Category

Figure 9 demonstrates the relationship between OPSG/TOI and Goals/TOI, Assists/TOI and Points/TOI using their team percentile within their position category. Figures 9a, 9b and 9c show forward results and Figures 9d, 9e and 9f show defensemen results. While these scoring metrics generally increase as OPSG/TOI rises, the results are more pronounced for forwards compared to defensemen. Team percentiles are presented to maintain player/team anonymity. It should be noted OPSG focuses on player movements while the puck is possessed; it does not evaluate a player's eventual decision with the puck.



(a) Forward Goals (b) Forward Assists (c) Forward Points



(d) Defensemen Goals (e) Defensemen Assists (f) Defensemen Points

Fig. 9. Relationship between OPSG/TOI and Scoring Statistics by Position Category

Table 1. Average OPSG/TOI and Correlation with OPSG/TOI by Position Category

Position Category	Avg. OPSG/TOI	Goals/TOI	Assists/TOI	Points/TOI
Forwards	0.0048	0.504	0.654	0.660
Defensemen	0.0017	0.419	0.344	0.363

Table 1 shows the average OPSG/TOI and the correlation between OPSG/TOI and each of the aforementioned scoring metrics by position category. OPSG/TOI's correlation with Goals/TOI, Assists/TOI, and Points/TOI is higher for forwards compared to defensemen. Because defensemen are usually last to lead the rush and position themselves from the point in the offensive zone, we hypothesize they are less likely to move the puck into a space that will generate dangerous opportunities for their teammates relative to forwards. As seen in Table 1, defensemen generate nearly three times less valuable space for their teammates compared to forwards. Thus, the relationship between OPSG/TOI and Assists/TOI is weaker than forwards with a correlation of 0.344 compared to 0.654. This follows intuition given that this metric is designed to measure how well players create space for others, as opposed to themselves, and forwards are consistently moving the puck within the offensive zone to create scoring opportunities for their teammates.

Additionally, the correlation for defensemen is higher in goals compared to assists. One reason this might occur is that offensive defensemen that generate goals are more likely to carry the puck in more valuable areas in the offensive zone. Considering there are two assists for each goal, defensemen do not need to be in these areas to generate assists. Deeper analysis is required to confirm the underlying cause for these relationships. We leave this for future work.

We can also analyze the OPSG/TOI breakdown by position, shown in Table 2. While left-wingers and centers seem to generate more OPSG/TOI compared to right-wingers, these are subject to the composition of the rosters analyzed, and thus, a more comprehensive analysis of this area is needed. A similar statement could be made on left and right defensemen.

Table 2. OPSG by Position

Position	OPSG/TOI
Right Wing	0.0044
Center	0.0048
Left Wing	0.0050
Left Defensemen	0.0018
Right Defensemen	0.0016

4.3 Team Evaluation

We calculate the OPSG Differential between the Home and Away team, along with their shot attempts, SOG, and goal differential for each game in the dataset. The correlations between these measures are shown in Table 3. OPSG has a stronger relationship with shot attempts and SOG compared to Goal Differential. As various pieces on expected goals have noted, goals occur more randomly in comparison to shot attempts and SOG [3, 5]. Exploring the slightly negative correlation between OPSG and Goal Differential is left for future work with a larger set of games.

Table 3. OPSG Differential Correlation

Metric (H)	Correlation
Shot Attempt Differential	0.647
SOG Differential	0.603
Goal Differential	-0.055

5 Limitations and Future Work

While the model we propose in this work helps characterize space creation, it has several limitations. A limitation of our Rink Control Model is that it is a descriptive model and is not calibrated to event outcomes e.g. passes/puck recoveries; thus the model may not be indicative of actual values of control beyond player orientation to a specified location. Additionally, our Rink Value Model assumes the receiver of a pass is able to shoot. In reality, the value of possession in different areas on the ice should incorporate all possible decisions available to a player and their expected outcome. With regards to our Transition Probability Model, we have not incorporated the NHL offside rule into our models; thus receivers may not be in positions to receive a pass despite having a high transition probability. Finally, our combined model does not value passing lanes differently from one another. While low-high and seam passes increase the probability of a shot being scored relative to other shots at the same location, our current model does not incorporate features to differentiate these types of plays. Each of these limitations is left for future work.

Furthermore, given OPSG results are limited to a 35-game sample, an immediate direction of future work is to run this process across a full season to better understand the consistency in a player's ability to create space with the puck and our metrics repeatability.

A puck carrier's teammates also have the responsibility of moving into positions where they can receive the puck and create scoring opportunities. To understand a player's off-puck space creation for teammates, we could follow the framework presented by Bornn and Fernandez [1] for Space Generation Gain (SGG). To understand a player's off-puck space creation for themselves, we could isolate a player's contribution to Transition Probability through their Pass Probability Fabric, their contribution to Rink Control through their individual influence as seen in Figure 1, and Rink Value which does not require any adjustments. We could then measure the changes in their probability of scoring on the next on-puck event assuming the possessor of the puck were to pass it to them. This could then be aggregated across all possessions.

Another direction of future work is to measure how well teams perform in blocking passing lanes using our pass probability model. Given a set of intended passes, we can calculate the probability of the pass being successful, and measure how teams block opponent passes compared to the probability of the pass being successful. This would require an accurate dataset of incomplete and blocked passes, and thus, we leave this study for future work.

An additional area of future work would be to better understand which types of players play best together, known as team formation problems in the broader artificial intelligence (AI) literature. Teams may not want to have three players who excel at on-puck space creation but are poor in off-puck space creation on the same line. These types of traits could be applied further into team formation algorithms to better predict team success.

6 Conclusion

The presence of spatiotemporal data in ice hockey shifts the types of questions that can be addressed through analytics and the methodologies employed to approach them. Research developed in other sports, most specifically soccer, lends itself well to ice hockey and allows for more complex approaches to these problems. In this paper, we present a new metric for ice hockey, On-Puck Space Generation (OPSG). This metric can be used to better understand play-making through a quantitative approach to space creation using Puck and Player Tracking Data. Our models can be applied at the game-level to understand which players are creating space and where it is most often being generated. Furthermore, we can learn which teams are best at collaborating to generate space for one another, and how this relates to collective success. We believe this framework represents the first step toward better understanding how players create space in valuable areas in ice hockey.

Acknowledgments. We would like to give special thanks to Dr. David Radke from the Chicago Blackhawks for his invaluable suggestions throughout this process. Additionally, we would like to thank Ryan Kruse, Sam Forstner, Mark Weinstein, Albert Lyu, Daniel Codos, Zack Brusso, and Chris McCorkle from the Chicago Blackhawks for their feedback. Finally, we would like to thank Brant Berglund and Keith Horstman at the NHL, as this project would not be possible without the use of PPT data.

References

1. Bornn, L., Fernandez, J.: Wide open spaces: A statistical technique for measuring space creation in professional soccer. MIT Sloan Sports Analytics Conference (2018)

2. Bornn, L., Fernandez, J., Cervone, D.: Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer. MIT Sloan Sports Analytics Conference (2019)
3. EvolvingHockey: A new expected goals model for predicting goals in the nhl (2018)
4. EvolvingHockey: Regularized adjusted plus-minus. https://evolving-hockey.com/glossary/regularized-adjusted-plus-minus/ (2021)
5. Green, S.: Assessing the performance of premier league goalscorers. Stats Perform (2020)
6. Lacey, P., Bialkowki, A., Carr, P., Matthews, I.: Quality vs quantity: Improved shot prediction in soccer using strategic features from spatiotemporal data. MIT Sloan Sports Analytics Research Paper Competition (2015)
7. Moneypuck: Shot prediction expected goals model
8. Radke, D., Brecht, T., Radke, D.: Analyzing passing metrics in ice hockey using puck and player tracking data. Linköping Hockey Analytics Conference (2022)
9. Ritchie, R., Harell, A., Shreeves, P.: Pass evaluation in women's olympic ice hockey. MMSports '22: Proceedings of the 5th International ACM Workshop on Multimedia Content Analysis in Sports (2022)
10. Spearman, W.: Beyond expected goals. MIT Sloan Sports Analytics Conference (2018)
11. Spearman, W., Basye, A., Dick, G., Hotovy, R., Pop, P.: Physics-based modeling of pass probabilities in soccer. MIT Sloan Sports Analytics Conference (2017)
12. Swartz, T., Wu, L.: A new metric for pitch control based on an intuitive motion model (2023)

Examining the Role of Hockey Leadership to Foster Inclusive Coaching Practices: Discussions from Atlantic Canada

Dr. Lynn LeVatte[1*], Dr. Christina Phillips[2], Shaun Ranni[3], Sarah MacRae[4], Dr. Kristin O'Rourke[1]

[1] Education Department, Cape Breton University, Nova Scotia, Canada
[2] Department of Curriculum, Teaching & Learning University of Toronto, Ontario, Canada
[3] Sport Leadership & Active Living Department, Cape Breton University, Nova Scotia, Canada
[4] Communications Department, Cape Breton University, Nova Scotia, Canada

Abstract. Coaching has been widely examined in the sport of ice hockey. Technical skill development, player management, and the ability to improve performance have been very notable areas of inquiry. As the critical roles of coaching leadership and communication become clearer, there is limited research available which explores the context of inclusive hockey coaching leadership to support more equitable practices. This paper will focus on specific data extracted from a previous study completed by the authors in which general hockey leadership skills and professional development were explored. This paper will present the outcomes of fostering inclusion and diversity from a coaching lens. Thirteen minor hockey coaches from Atlantic Canada (i.e., who are members of the Atlantic Hockey Group) participated in this qualitative study. Semi structured interviews were conducted online or in-person. A thematic analysis was used to explore data obtained from the interviews. Results revealed that coaches had limited communication training experience when working with diverse abilities, age groups, languages, genders, or cultures. Limited professional development specific to inclusive training was noted by participants. Our results demonstrated that various self-led leadership strategies were utilized to promote inclusive practices such as informal community-peer mentorship opportunities, and small group instructional sessions. Overall, the results give us insights into coaches' experiences with inclusive leadership and highlight current gaps. During the conclusion, future recommendations for continued study, specifically within leadership training for diversity within ice hockey, are offered.

Keywords: Communication, Ice Hockey, Coaching Leadership, Cultural Diversity, Inclusion, Performance, Player Engagement

* e-mails: Lynn_levatte@cbu.ca, c.phillips@mail.utoronto.ca, Shaun_ranni@cbu.ca, Sarah_macrae@cbu.ca, Kristin_o'rourke@cbu.ca

1 Introduction

1.1 Inclusivity in Sports

Sports are an integral part of an active and healthy lifestyle (Statistics Canada 2023). According to Statistics Canada (2023), Canada is a leading sport nation. There are several factors which contribute to participation, such as the season and the geographic and social diversity in the area. Hockey, our national winter sport, was invented in Canada in the 1800s, and basketball was invented by Canadian Dr. James Naismith in 1891 to condition young athletes during the winter. Other sports, such as soccer and basketball, are also popular. According to Gough (2023), hockey is the most popular sport in Canada, followed by soccer and basketball.

The influence of coaching behaviors on athlete performance has been one of the most investigated topics in coaching science (Gilbert & Trudel, 2004). Coaches are engaged with providing and assessing the development of various technical skills and monitoring training at various levels. Performance assessment has become a field of expertise that is crucial for researchers and practitioners (e.g., coaches, strength and conditioning coaches, scouts, program directors), who need to be aware of the mechanisms that predispose hockey players to perform in key situations (Bournival et al., 2023).

The National Hockey League (NHL) is the leader for many directives within ice hockey. Understanding communication, coaching leadership and cultural diversity has been examined recently. NHL commissioner Gary Bettman has stated that, "we are working to better understand and accelerate our engagement across all layers of diversity, including nationality, race, gender identify, sexual orientation, disability and religion," (Wyshnyski, 2022). We know very little about hockey coaching communication within the realms of inclusion and diversity. More research is needed to better understand perceptions, best practices, and inclusive practices to employ within ice hockey coaching specifically to improve coach engagement, player experiences, and hockey performances. Supporting this development, NHL commissioner Bettman stated that "Each day, we are committed to ensuring inclusion becomes more of 'who we are' than 'what we do" (Wyshynski, 2022).

According to Oldham (2022), society must break down the interlocking forms of discrimination and social injustices at the junior, college, and professional levels of ice hockey. Furthering a notion of "Why does this happen? "and "How can we fix it?" Leadership in sport is an emergent field which has been gaining attention on a national level (Jones & Khan 2017). More specifically, understanding the optimal leadership training for more inclusive hockey coaching is intriguing. The objective of this paper is to document and examine the communication experiences of minor hockey coaches in Canada who support youth with diverse needs and determining what resources are required to improve coaching experiences. According to Matthews and Erickson (2023), youth sport is a context that can promote positive youth development (PYD), with coaches being a key agent for positive development. Furthermore, transformational leaders employ four strategies, colloquially referred to as idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration (Matthews & Erickson, 2023, p.1). Lara-Bercial & Mallett, (2016) investigated characteristics of coaches and their relationship to leadership. The findings of this study revealed that coaches had a common set of personal characteristics, which included an exceptional work ethic, strong communication skills, a quest for continuous improvement, and effective leadership behaviors that inspired their athletes. This research study will explore coaching leadership through an inclusive lens. As described by Liew and McTigue (2010), educating the "whole child" became more prominent and thus enhanced the teaching and coaching skills of professionals who work directly with youth. Leadership and training are imperative to properly address the expansive growth and popularity of Canadian hockey.

1.2 Hockey and Canada

As the IIHF reported, there are currently 513,684 Canadian hockey players registered (IIHF, 2024). The total number of players registered with the IIHF is 1,563,749, meaning that Canadian players account for 1/3 of the total membership of an organization that governs 81 countries. Hockey is considered a cultural truism and a way of life among Canadians, with a connection so powerful and strong that it has united a vast nation from coast to coast (Cairnie, 2019). In fact, the sport is often considered a Canadian national treasure for its ability to build kinship bonds between ethnicities, classes, and cultural groups. While such perceptions of inclusivity have remained prevalent in the sport, recent critical events, such as abuse scandals and racism, have negatively impacted the sport and its leaders (Burke, 2022).

1.3 Inclusive Hockey Leadership & Coaching

From an understanding of the early vision that guided Canadian sport history, its is readily observable that promoting inclusivity in Canadian sport has been recognized as critically important and can be best described as follows:

> "Sport is welcoming and inclusive, offering an opportunity to participate without regard to age, gender, race, language, sexual orientation, disability, geography, or economic circumstances" (Canadian Sport Policy, 2002, p.13).

In December 2021, the Prime Minister of Canada released a mandate letter, providing clear direction on the importance of incorporating the views of Canadians when considering society, inclusivity, diversity, and our historically marginalized communities. As he noted:

> "We must continue to address the profound systemic inequities and disparities that remain present in the core fabric of our society, including our core institutions... I expect you to include and collaborate with various communities, and actively seek out and incorporate in your work,

> the diverse views of Canadians. This includes women, Indigenous Peoples, Black and racialized Canadians, newcomers, faith-based communities, persons with disabilities, LGBTQ2 Canadians, and, in both official languages." (Kay et al., 2022).

Providing equal opportunity and accessibility is an imperative need within Canadian sport communities. As noted by Kay et al. (2023), sport research findings suggest that research participants felt efforts should be made to increase the participation of under-represented groups in sport. Particularly, these groups included indigenous people, racialized people, women and girls, persons with disabilities, children and youth, new Canadians, and economically disadvantaged people. Hockey Canada, the governing body of amateur hockey in Canada addressed this need recently. In October 2020, the Equity, Diversity and Inclusion (EDI) task team provided a report to the Hockey Canada Board of Directors that included a framework for the strategic plan on EDI. Additionally, organizations require strategies to support the successful engagement of hockey players from these underrepresented groups (Hockey Canada, 2023). Hockey Canada has published its first Equity, Diversity, and Inclusion (EDI) Path Forward, which includes a Commitment to Action statement that summarizes the organization's ongoing work to drive long-term, sustainable change within the hockey ecosystem in Canada, building an environment where people feel valued for their differences and have positive experiences with hockey (Hockey Canada, 2023).

According to Matthews and Erickson (2023), youth sports can promote positive youth development (PYD); coaches are a key agent for positive development. Furthermore, transformational leaders employ four strategies, known as idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration (Matthews & Erickson, 2023, p.1). Duguay et al. (2020) found that coaches who embody effective leadership qualities not only impart essential athletic skills but also instill crucial life lessons, such as teamwork, discipline, and resilience. Additionally, inspirational leaders serve as role models, encouraging young athletes to strive for excellence while maintaining a positive attitude, sportspersonship, and respect for others (Davis, 2018). Conversely, poor leadership can have detrimental effects, leading to a toxic atmosphere, demotivation, and even dropout rates among youth participants (Fouraki et al., 2020). Fostering an environment of belonging and inclusivity is essential for all youth hockey players and may be difficult for coaches. Jedwab & Holly (2021), researched immigration and diversity in ice hockey and concluded that despite the many challenges hockey faces as the country's most watched and played sport, it still creates a strong sense of community belonging amongst those who choose to play (p.164). Providing coaches with appropriate training and leadership skills to support diversity is critical (Saotome, 2013). Additional research concludes that personality, character, and communication skills are essential (Erickson, 2023). In addition to coaching leadership research, evidence suggests that youth with strong leadership skills are more likely to have positive work and family relationships, enter and graduate college, succeed in their careers, and have better mental and physical health outcomes (Greenberg et al., 2017). Previous research has determined that when enhancing learning, one important factor is the ability and experiences educators and coaches possess to engage and support learners (Jones & Kahn, 2017).

2 Methodology

This qualitative research study utilized open-ended, semi structured interview questions to collect data specific to the training and coaching experiences of hockey coaches from Atlantic Canada. Data were collected for this study from November 2022-May 2023. Participants were recruited by email from the project partner, The Atlantic Hockey Group (AHG). The participant sample (n=13) comprised of minor hockey coaches who volunteered with the AHG and instructed youth aged 4-18 years. The sample for this research was purposive, as ice hockey coaches, both male and female, were invited to participate. Recruitment also involved notices through social media and direct email from the Atlantic Hockey Group. There were small inducements of Tim Horton coffee cards offered for study participation.

Participation was voluntary, and coaches were invited to complete a short interview with a member from the research team. The interviews took place in-person and/or online, via Zoom TM depending upon location and availability of each participant. All participants self-identified as male (i.e., n=13) who ranged in age from 23 to 52 years. The average span of coaching experience was 9 years.

The qualitative interviews were composed of five open-ended questions. For this paper, three questions served as focal points of analysis. Interview questions analyzed specifically for this paper included:

1. Social-Emotional Learning includes aspects of enhanced leadership, empathy, understanding, self-regulation, behavior support, trust, honesty, inclusivity, etc. What is your experience with these specific components? Were they taught explicitly or included within your coaching training? If yes, what types or when?
2. In your coaching career, how often are you provided with leadership training? What types of training did you receive as professional development?
3. When coaching young ice hockey players, what is the most challenging aspect in terms of connections and relationship building with your players and/or families? Are there other barriers or challenges with your players? What types of training do you feel would be beneficial for coaches?

Researchers took field notes at the end of sessions to ensure key messages were highlighted. Sessions were also recorded with permission for transcription. Ethics approval was obtained from Cape Breton University prior to engaging in our interview process.

Research Question: What are the hockey leadership experiences of coaches in Atlantic Canada?

2.1 Theoretical Framework

This study was framed upon aspects of Social Emotional Learning (SEL). Encompassing approaches where youth and children learn to recognize and manage emotions, develop positive relationships, behave ethically, care about others responsibly, make good decisions, and avoid negative behaviors (Gould et al., 2022). It involves "…teaching children to be self-aware, socially cognizant, able to make responsible decisions, and competent in self-management and relationship skills…" (Zins et al., 2007, p. 195). Social Emotional Learning (SEL) is critical for children and youth's long-term success in and out of school (Weissburg et al., 2015). Examining the intrapersonal characteristics for success, in-depth personal reflection, emotional intelligence, and a quest for continuous improvement have been instrumental within SEL research (Domitrovich et al., 2017).

SEL can be used to promote character development among athletes (Elias, 2016). The term "Educational Athletics" is used by the Massachusetts Interscholastic Athletic Association to express how athletics and competition can be used as an extension of the classroom and an educational activity to teach life lessons and prepare young people with values for lifelong learning (Elias, 2016).

3 Findings

In this section, one major, overarching theme will be presented through the analysis of three questions posed to hockey coaches, reflecting experiences within a leadership and inclusive lens. These semi-structured interviews were transcribed verbatim and filed within a MicrosoftTM Office Teams account. The first author and researcher read and reviewed the raw data transcripts sets several times and listened to audio files during the analysis to assist with the conceptualization of ideas presented. Data were organized and analyzed using a coding process that led to the construction of themes (Saldana, 2014). Inductive content analysis was employed as this project included non-complex research and the sample size (n=13) was small (Vears & Gillam, 2022). An inclusive theme among the research team emerged during our iterative, co-constructed analysis processes. As noted by Williams and Moser (2019), coding in qualitative research comprises processes that enable collected data to be assembled, categorized, and thematically sorted, providing an organized platform for the construction and development of meaning. In this article, one main theme is presented, highlighting inclusive practices for hockey leadership.

3.1 Theme 1: Development for Inclusive Leadership Practice

A key theme of inclusive leadership emerged as a result of data analysis from interviews with 13 minor hockey coaches. Participants in this study expressed the need for leadership training and development specific to coaching youth with varying hockey skills, cultural backgrounds, gender, and language capabilities. While participants attempted to recall types of professional development training they had received, many found this task very difficult.

Gender. In some cases, participants described challenges when coaching self-identified males and females on the same team. They felt a disconnect in team cohesion due to having separate dressing rooms for males and females, which left the female ice hockey players feeling isolated. Participant 2 illustrated this complexity in stating,

> ... I have a female hockey group, I am a male coach and its different than coaching males... we work on lots of skills and I need to know what is happening in their lives because if they are having a bad day at school or failed a math test they will show it... I have to be careful not to push them to far, right? ... I do that on my own; it's important to build relationships...

Participant 10 also addressed gender in hockey and the complexities regarding leadership training in this quotation,

> ... there's no behavioral, you know, teaching classes or any kind of courses like that let's put it that way. So that's really not at all. It's a very, you know, male dominated industry, I would say... push to have, you know, girls involved in coaching positions, things like that so like the diversity aspect... and you know welcoming everyone has changed...

Relationship and Leadership Building. Participant 1 highlighted the importance of building relationships with parents,

> To get to know the parents off the ice. Because everyone's schedules are so busy and hockey practices/ games are at weird times. Getting to know them more is the most challenging part. The hockey tournaments are beneficial because at a hotel the parents get can meet each other. During the weekday everyone's got to get home and get ready for the next day. As a coach you're the first one there and the last one out so mostly everyone is gone by the time you're out...

Informal community peer mentorship was identified as crucial for understanding and learning about leadership. Participant 8 also noted that the positive and welcomed impact of community mentors. He stated that,

> "... the hockey code mentor was pretty good. I had a guy, who is at Dalhousie, hockey team, so he gave me some drill ideas if I asked, and we walked a couple of our practices, and he came out a couple times..."

Participants often shared how they created relationships with community partners and other coaches. Peer mentoring was described as informative and useful, providing an informal opportunity for less experienced coaches to learn from more experienced or diversified coaches. Participant 3 shared his personal experience with local coaches. He stated that,

> ... the Atlantic Hockey Group team that I did, I was basically the head coach and manager of. I didn't really, I did everything myself. Basically,

> I took a couple of my buddies on to help me out when it came to like the on- ice stuff, just to be around and push some pucks. And if they had any drills, I made sure that the guys that I had on the bench were hockey guys too…

Further, participant 2 explained that communicating expectations is vital for success:

> "I think one thing is coaches need to be exposed to understanding team dynamics and how those things happen. So, there's, yes, you have expectations of the team. There are overarching expectations".

Participant 1 explained that,

> …There's been a whole new level of respect and peer coaching that we didn't have before. So, this year I got another player new to my team whose brother is actually the captain of the same major bantam team. What do we need to work on? And plan practices around their feedback and then really have the kids learn from their peers. They may have more talent on the ice, but they're kids that will not socialize. They sit in the corner in the dressing room. They won't talk to anybody else. They you know, you go away for a tournament, they won't go out to supper. They'll stay in their room. It is imperative to appreciate the need and desire for minor hockey coaches to identify existing gaps within leadership training. It is from this identification that professional development can be employed where and when necessary…

Participant 3 summarize his thoughts on mentorship and its value:

> "My first year that I got involved, I actually was handed head coach of the skills group, didn't even get to assistant, didn't even get to help out for a year. I was handed it the first year basically. Very good. I'll call it Mentor ahead of me that handed over all his notes and took the coaching clinics from hockey."

In alignment with the other participants, participant 2 reported:

> …this is the gap right now within our hockey communities. Is that there's a lack of understanding that we're building leaders…so what we say is all we're focused on the kids, and we forget that we're also developing coaching staff and future leaders in that… but the aspect of those other pieces of growth and leadership is where we could really improve, and we could build on…

Supporting Players with Disabilities. Describing some of the most challenging aspects of coaching a player with diversity and disabilities, participant 1 noted that,

> ...some of those Hockey Canada modules, we did a lot on diversity for sure. Biggest training would be working with kids who have like ADHD and stuff like that it kind of it kind of misses that number of years we've had 4-5 maybe six kids that if levels of ADHD and I wish I was more training I guess on that part of things sure like diversity they cover very well yeah but I find kids with ADHD there needs to be more training...

The additional need for training to support successful coaching was also highlighted by participant 1 as noted here,

> "...I received ...not a whole lot other than what's covered in hockey Canada that's mandatory. We would have to volunteer to take on others on our own time. I wish there was a lot more..."

Conversely, participant 4 recalled discrimination in sport and fair play course trainings that were offered through Hockey Canada and this training was augmented with additional courses added by Atlantic Hockey Group.

> "That is one thing I can say, like to list couldn't possibly remember the courses, but they are great to make sure the coaches have what they need". Participant 9 expressed the need for additional and frequent training: "I'd say professional development. I haven't really seen any other than the requirement, which is every three years to do the courses."

Language and Cultural Diversity. During the interviews, participants discussed important experiences coaching culturally diverse teams. For example, participant 3 revealed the following:

> ...but I think when it comes to like the inclusivity part, I think like we've had over the years, we've had a lot of kids from Indigenous communities and ask us only playing on our team... compared to kids who are from Non Indigenous communities, who are not really around that much unless they're playing against them... to bring them in on a team and have them playing together, it was kind of like a different thing for the kids out in the Bay where there's not many First Nations people out there...

Communicating with youth and parents/caregivers who have language or cultural diversities can be challenging for coaches. The participants agreed that when coaching a player with language barriers, this can be difficult to engage the player within the team game plans and practices or it can also hinder the team bonding. As participant 4 revealed,

> Uh, sometimes it can be just a simple thing like language. Like, I coach quite a few kids from all over the province. Yeah, you know, so like, a lot of times you do, you have like, your Francophones and your French speaking stuff. This can be difficult, right ?...but, nothing... Can't be overcome.

The training gap was identified again by participant 8 who explained,

> "...and, I believe there's one coaching... development at the beginning of the year, um, but other than throughout the year, it's, if there's no complaints, then nobody really says anything to you."

Participant 10 shared a valuable case, specific to cultural diversity where he stated that,

> I would like to see maybe more education along the way on how to treat people. Times have changed, you know diversity and inclusion is a big part of you know the sport that really there's not a lot of training about, you know, so I would like to see some stuff like that had, how did you come from Ukraine? We've got, you know, roles and different ethnicities

In conclusion, some participants noted that they felt unprepared to coach players with diverse cultural backgrounds and language.

4 Discussion

The findings of this study emphasize the importance of current professional development training required for the changing landscape of ice hockey in Canada. Coaches are essential to hockey at every level. A caring, enthusiastic, well-trained coach can positively influence the experiences of players, parents, and other coaches (Hockey Canada, 2024). Provincial hockey coaches receive certification from the National Coaching Certification Program (NCPP). Hockey Canada works with local hockey associations across Canada to provide effective education, certification, and registration for thousands of hockey coaches annually. The Hockey Canada Learning Lab has also been recently launched to provide additional coaching training resources (Hockey Canada, 2024). To remain active within the NCCP program, hockey coaches must maintain their status by obtaining professional development (PD) points. Activities such as e-learning hockey modules workshops, judging, facilitation, committee work, and active are all accepted PD points (Hockey Canada, 2024).

Communication, inclusive practices, and the need for additional training and professional development were focal points addressed by participants. This study suggests that many coaches had received technical training in ice hockey coaching; however, there is a need for additional training that addresses parent communication and coaching for diverse populations and cultures. Coaches also described the distinct difference between community and formal mentorship training and mandated technical development training. Previous reports have suggested that professional development training in these aspects can be beneficial in sport coaching (Shen, Rose & Dyson, 2022). This research may inform coaches, hockey administrators, players, and parents about inclusive hockey practices that can support gender diversity, players with disabilities, and cultural awareness.

Important findings emerged from discussions surrounding the need for accurate and frequent development training that included gender, cultural diversity, and communication best practices. Implications from this study support the ongoing need for training and professional development, addressing existing gaps in current training, and the changing culture within ice hockey from a male-dominated sport to one more welcoming for all abilities and all players.

Suggestions for future study include additional research in the area of ice hockey coaching, particularly focusing on diversity and inclusive application. Limitations and research directions for qualitative inquiry noted in the study included sampling from only one hockey organization. While the present study explored coaching leadership in ice hockey and the aspects of inclusivity, cultural diversity, and gender, the findings and discussion focused on the participant's experiences in coaching. All participants had coached hockey for the AHG in Atlantic Canada. Additional and continued research is also needed to explore training delivery modalities and specific content for leadership and professional development. Future qualitative studies may also focus specifically on these inclusive sub themes to cultivate an in-depth understanding of coaching leadership.

Replicating the study within other regions of Canada would also be important to allow for a larger sample and broader base of study. Based on the findings of this research, the authors recommend a streamlined access point or tab system that directly outlines PD links from the Hockey Canada website, as accessing the e-training modules was difficult. Data from this study revealed several challenges in coaching players with disabilities or language barriers. Offering PD training such as Inclusive Coaching Practices for Players with Intellectual Disabilities may be helpful. New courses supporting effective communication practice with adolescents may also be beneficial. Community coaching mentorships with experienced junior-level coaches could be recommended to support successful leadership. Engaging local minor hockey associations with local Junior level coaches could potentially provide practical and experiential learning opportunities.

5 Conclusion

This study explored leadership and professional development training experiences among hockey coaches in Atlantic Canada. Specific findings examined inclusive practices within hockey coaching. These emerged as topic areas of gender, relationship, and leadership building, supporting players with disabilities, and players with cultural or language diversities. The results suggest that most coaching training is done by Hockey Canada and provincial associations, and informal training was provided through local community hockey teams or via peer coaches. Additional training that focuses on inclusion, diversity, adolescent development, and communication within hockey realm is necessary. Hockey culture is evolving, and coaching needs to support a much more diverse population.

6 Acknowledgements

We would like to thank the Atlantic Hockey Group for its ongoing support during this study, especially regarding the recruitment of participants.

References

Adams, C. & Leavitt, S.: It's just girls' hockey: Troubling progress narratives in girls' and women's sport. International review for the sociology of sport 53.2, 152-172 (2018)

Anderson-Butcher, D., Bates, S., Amorose, A. et al.: Social-emotional learning interventions in youth sport: what matters in design? Child and Adolescent Social Work Journal 38.4, 367-379 (2021)

Bournival, M., Martini, G., Trudeau, F., & Lemoyne, J.: The science and art of testing in ice hockey: a systematic review of twenty years of research. Frontiers in Sports and Active Living, 5, 1252093–1252093. `https://doi.org/10.3389/fspor.2023.1252093` (2023)

Camiré, M.: Benefits, pressures, and challenges of leadership and captaincy in the National Hockey League. Journal of Clinical Sport Psychology 10.2, 118-136 (2016)

Camiré, M., Trudel, P and Bernard, D. A case study of a high school sport program designed to teach athletes life skills and values. The Sport Psychologist 27.2, 188-200 (2013)

Cairnie, J.: Truth and reconciliation in postcolonial hockey masculinities. Canadian Literature 237, 103-119 (2019)

Domitrovich, C.E., et al.: Social-emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. Child development 88.2, 408-416 (2017)

Elias, M.: 4 ways to use athletics to promote social emotional learning and character development. https://www.edutopia.org/blog/4-ways-use-athletics-promote-social-emotional-learning-and-character-development-maurice-elias. (2016)

Statista, Most popular sports in term of interest in Canada. https://www.statista.com/statistics/1447847/favorite-sports-interest-canada/#statisticContainer

Gould., D. Martin, E., Walker, L.: A season-long investigation of social emotional learning associated with high school basketball participation, Journal of Applied Sport Psychology, 34:6, 1102-1124, `https://doi.org/10.1080/10413200.2021.1955421`. (2022)

Hockey Canada: Statistics & history. Available at: http://www.hockeycanada.ca/en-ca/Hockey-Programs/Female/Statistics-History.aspx last accessed May 2023

Hockey Canada, Hockey Canada published equity, diversity, and inclusion path forward report. https://www.hockeycanada.ca/en-ca/news/edi-path-forward-published-2023-corp.

Hockey Canada, Equity, diversity, and inclusion path forward. https://cdn.hockeycanada.ca/hockey-canada/Corporate/About/Downloads/2023-edi-path-forward-e.pdf

Hockey Canada, Maintaining your coaching certification. https://www.hockeycanada.ca/en-ca/hockey-programs/coaching/essentials/ nccp/maintain-certification.

International Ice Hockey Federation, Survey of players. http://webarchive.iihf.com/iihf-home/the-iihf/survey-of-players/index.html last accessed 2023/05/10

Jedwab, J., & Holley, P.: Who's the north? the challenge that immigration and diversity present to the dominance of hockey in 21st century Canada. Canadian Ethnic Studies, 53(3), 163-181 (2021)

Jones, S. & Kahn, J.: The Evidence Base for How We Learn: Supporting Students' Social, Emotional, and Academic Development. Consensus Statements of Evidence from the Council of Distinguished Scientists." Aspen Institute (2017)

Kay, J., Larue, L. Steele, K., Gendron, B.: Shaping inclusive sport policy and programs. https://sirc.ca/blog/shaping-inclusive-sport-policy-and-programs/.

Lara-Bercial, S. & Mallett, C.: The practices and developmental pathways of professional and Olympic serial winning coaches. International Sport Coaching Journal 3.3, 221-239 (2016)

Liew, J. & McTigue, E. Educating the whole child: The role of social and emotional development in achievement and school success. Handbook of Curriculum Development (2010)

Liew, J., Cameron, C. E., & Lockman, J. J.: Parts of the whole: Motor and behavioral skills in self-regulation and schooling outcomes. Early Education and Development, 29(7), 909–913. https://doi.org/10.1080/10409289.2018.1500513 (2018)

Luna, P., Guerrero, J., & Cejudo, J. (2019). Improving adolescents' subjective well-being, trait emotional intelligence and social anxiety through a programme based on the sport education model." International Journal of Environmental Research and Public Health 16.10, 1821 (2019)

Shen, Y., Rose, S. & Dyson, B. (2022) Social and emotional learning for underserved children through a sports-based youth development program grounded in teaching personal and social responsibility, Physical Education and Sport Pedagogy, https://doi.org/10.1080/17408989.2022.2039614 (2022)

Sport Information Resource Centre, Canadian Sport Policy.https://sirc.ca/wp-content/uploads/2019/12/2002-the_canadian_sport_policy.pdf.

Statistics Canada, Sport in Canada. https://www.canada.ca/en/canadian-heritage/services/sport-canada.html.

Urquhart, D., Bloom, G. A., & Loughead, T. M.: The development, articulation, and implementation of a coaching vision of multiple championship–winning university ice hockey coaches. International Sport Coaching Journal, 7(3), 335–346. https://doi.org/10.1123/iscj.2019-0096 (2020)

Vears, D. & Gillam, L.: Inductive content analysis: A guide for beginning qualitative researchers. Focus on Health Professional Education, 23(1), 111–127. https://doi.org/10.11157/fohpe.v23i1.544 (2022)

Weissberg, R.P., Durlak, J.A., Domitrovich, C.E., & Gullotta, T.P.: Social and emotional learning: Past, present, and future. In J.A. Durlak, C.E. Domitrovich, R.P. Weissberg, & T.P. Gullotta (Eds.), Handbook of social and emotional learning: Research and Practice, pp. 3-19, Guilford Press: New York (2015)

Williams, M. & Mosher, J.: The art of coding and thematic exploration in qualitative research. International Management Review, 15, 45, (2019).

Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H. J. The scientific base linking social and emotional learning to school success. Journal of Educational and Psychological Consultation, 17(2–3), 191–210. https://doi.org/10.1080/10474410701413145 (2007)

Characterizing Playing Styles for Ice Hockey Players

Anton Olivestam[1], Axel Rosendahl[1], Erik Wilderoth[2],
Niklas Carlsson[1], and Patrick Lambrix[1]

[1] Linköping University, Sweden
[2] Färjestad BK, Sweden

Abstract. Although analytics is being used in, e.g., the evaluation of players and scouting, it is still challenging to quantify skills and playing styles of players. Such information is important for roster creation and scouting, where teams want to find players that have a playing style that fits within the team, as well as for game preparation to understand the playing style of opponents. In this paper we use player vectors to characterize a player's playing style. The player vectors contain representations of skills that are computed from game event data. Further, we use fuzzy clustering on the vectors to generate five types of defender playing styles and five types of forward playing styles. For these types, we show the typical skill levels and players with similar styles.[1]

1 Introduction

In ice hockey, the general manager and scouts are responsible for identifying the most skilled players to build a high-performing team within their budget limits. Historically, the teams relied on the manual analysis conducted by scouts and general managers. The introduction of data analytics transformed the process for scouting hockey players [18]. By leveraging data-based metrics, teams were able to adopt a more objective approach to decision-making, particularly in evaluating player performance. However, despite the growing influence of analytics in areas such as scouting and player performance evaluation, quantifying the nuanced skills and playing styles of individual players remains a challenge. Such information is important for roster creation, where teams want to find players that have a playing style that fits within the team.

In the complex and dynamic game of hockey, characterizing playing styles is challenging due to multiple aspects. One such aspect is that certain features are difficult to quantify, making it a challenge to identify appropriate variables for model building. Moreover, skills are typically not explicitly encoded and stored in the available event data, but rather must be derived from player actions or performance across various event types. For instance, a certain kind of defensive skill may depend on how well the defender performs in different aspects such as blocked shots, hits, and dump outs.

---
[1] Equal contribution for the first two authors.

In this paper, we characterize the playing style of a player in a data-driven manner. We use event data from hockey games of three leagues and 2.5 seasons (Sect. 3) to define different kinds of skills for defenders and forwards (Sect. 4). A player's playing style is then represented by how the player performs for these different skills. Formally, the player's playing style is represented by a skill vector. Further, we use fuzzy clustering to define five distinct playing styles for both defenders and forwards (Sect. 5). We show the typical skill levels for prototypical players in the clusters and give examples. Information about the clusters can be used for scouting, e.g., for finding players with similar playing style, as well as for game preparation, as understanding the playing style of the opponent can offer certain tactical advantages.

2 Related Work

There has been research on different topics that relate to trying to characterize and compare players in ice hockey.

One approach has been to define performance metrics. Many performance metrics assign values based on particular types of actions in the game. For instance, goals, assists, and Corsi attribute a value to goal-scoring actions, to passes that lead to goals and to different types of shots, respectively. Variants of traditional metrics have been proposed such as regression models replacing the +/- measure (e.g., [16,17,7]). In [8] principal component analysis was performed based on 18 traditional metrics and a performance metric based on the four most important components was proposed. More recent work takes the context in which actions are performed into account. For instance, [20] attributes value to goals, but the value of the goal depends on the situation in which it is scored. Event impacts for different kinds of actions in [26] are based on the probability that the event leads to a goal (for or against) in the next 20 seconds. Several works model the dynamics of an ice hockey game using Markov games (e.g., [29,9]). In [21,27,28,13,24,15] action-value Q-functions are learned with respect to different targets. Different goal-based performance metrics taking the importance of goals into account are defined in [10,22]. Player rankings are presented in [25,14,11].

Another approach uses a probabilistic method for quantifying player roles in ice hockey. Earlier work allowed for a player to only belong to one role [30,2], while more recent work allows for a player to belong to different roles with some probability [23]. In the latter case, players can be compared based on their membership in different roles.

In other sports player vectors have been used to try to characterize a player's playing style. In [6], a basketball player's defensive play is characterized by shot taker, shot location, and expected outcome of the shot. In [3], a football player's playing style vector represents the areas on the field where the player tends to be and which actions in terms of passes, dribbles, crosses, and shots, the player performs in these areas. Movement patterns in shot situations in football were predicted in [12].

3 Data

The data we used is a proprietary dataset produced by Sportlogiq [2]. The dataset consists of event data from the following leagues: Swedish Hockey League (SHL), Hockeyallsvenskan (HA) and the American Hockey League (AHL). The choice of these three leagues originates from the fact that many transfers happen between these leagues. For instance, traditionally, many imports to the SHL come from the AHL, and many drafted SHL players start out in the AHL. As HA is the league one step down from SHL, SHL is a natural next step for many HA players. The data includes complete seasons for the three leagues for 2021/2022 and 2022/2023, as well as data from the 2023/2024 season up until January 28th, 2024. In total the dataset contains 7,532 games, 4,014 unique players, 68 unique teams and 28.5 million events. An event is described by more than 50 different parameters. The dataset comprises 2,553 forwards, 1,393 defenders, and 452 goaltenders, totaling 4,398 players. This figure exceeds the number of unique players, capturing that some individuals have played both forward and defender.

4 Player Vectors

4.1 Feature selection

Based on domain expert knowledge, we decided to use the skill sets as shown in Tables 1 and 2, for defenders and forwards respectively. For each skill, we utilized dataset features that influence it. Examples are given in the tables.Features can belong to different skills. For defenders there are 13 different skills that are described by five to seven features/actions and for forwards there are 18 different skills that are described by two to seven features/actions.

Table 1: Skills and example actions for defenders.

Skills	**Actions**
Passing	e.g., different types of passes
Skating	e.g., exits, entries, dumps
Shooting	e.g., different types of shots
Defensive Stickwork	e.g., blocked passes, loose puck recoveries
Puck Moving	e.g, some types of passes, dump-in recoveries
Point Producing	e.g., different offensive zone events
Powerplay Playmaking	e.g., powerplay playmaking events
Powerplay Scoring	e.g., powerplay shots and goals
Physical Play	e.g., body checks and defensive plays
Slot Defense	e.g., blocked shots and dump outs
Stay at Home	e.g., different defensive zone events
Penalty Killing	e.g., different penalty killing events related to puck recovery
Penalty Killing Slot Defense	e.g., different penalty killing defensive plays

[2] `https://www.sportlogiq.com/hockey/`

Table 2: Skills and example actions for forwards.

Skills	Actions
Passing	e.g., different types of passes
Skating	e.g., different types of controlled entries
Powerplay Playmaking	e.g., different types of controlled entries and passes in powerplay
Powerplay Slot Engagement	e.g., powerplay actions close to net
Powerplay Scoring	e.g., powerplay shots and goals
Defensive Puck Control	e.g., dump outs and loose puck recoveries
Defensive Zone Play	e.g., different defensive zone actions
Defensive Positioning	e.g., blocked shots and passes
Slot Defense	e.g., rebounds and dump outs
Penalty Killing	e.g., shorthanded defensive plays
Slot Engagement	e.g., offensive actions close to net
Heavy Game	e.g., body checks and defensive plays
Forechecking	e.g., offensive zone loose puck recoveries
Cycling the Puck	e.g., puck protections and receptions
Neutral Zone	e.g., different neutral zone actions
Puck Moving	e.g, some types of passes, entries
Offensive Zone Play	e.g., different offensive zone events
Shooting	e.g., different types of shots

4.2 Player vector construction

The player vectors are constructed based on the skills in Tables 1 and 2. First, all players who played less than 200 minutes were filtered out. Next, 13 feature vectors were created for each defender and 18 feature vectors for each forward. Each of these feature vectors quantifies a skill and contains the frequency of each action that describes that skill. For instance, for a particular defender, a defender skill with seven actions is represented by a vector of length seven where each element in the vector represents an action and its frequency for the defender in the dataset.

After constructing all feature vectors, we normalize them based on the player's ice time, i.e., the values are divided by the time on ice (TOI) of the player and multiplied by 60 (where 60 minutes is the length of a game in regulation time). This ice-time normalization was done to address potential differences in event frequency attributed to playing time disparities. Further, the values are standardized using MinMaxScaler in the scikit-learn library for Python [19]. This method transforms each feature by scaling it to a value between 0 and 1. This was done to take into account that some events are more frequent than others in a game and would otherwise have an undesired larger weight. For example, a pass happens much more often than a body check. In Fig. 1 we show the distribution of some events in the dataset. We note that passes, pass receptions, and loose puck recoveries account for the majority of events in the data with a total of 65.1% of all events.

Further, we apply non-negative matrix factorization (NMF) to each feature vector using the NMF in the scikit-learn library for python [19] to reduce its dimensionality to a single component. Thus, after this operation, every skill is

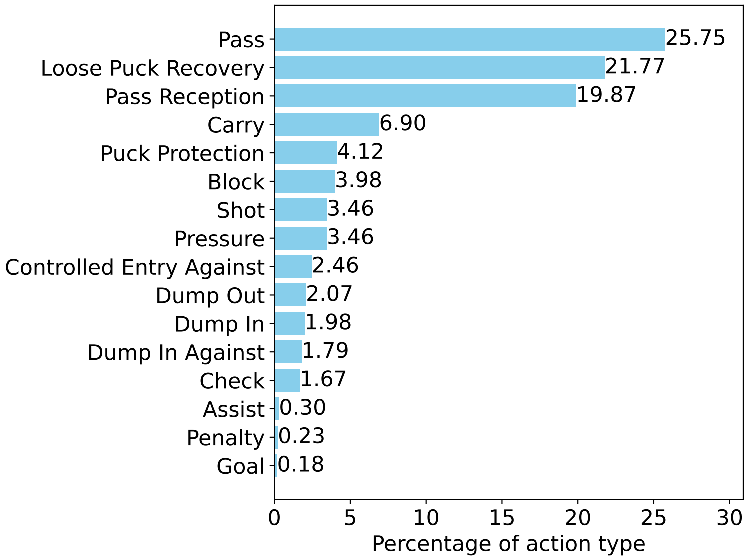


Fig. 1: Distribution on event frequency.

represented by one feature. Note that this operation leads to the fact that some values are higher than 1. Figs. 3a and 4a show the distributions of the values for the skills for defenders and forwards, respectively, using boxplots[3].

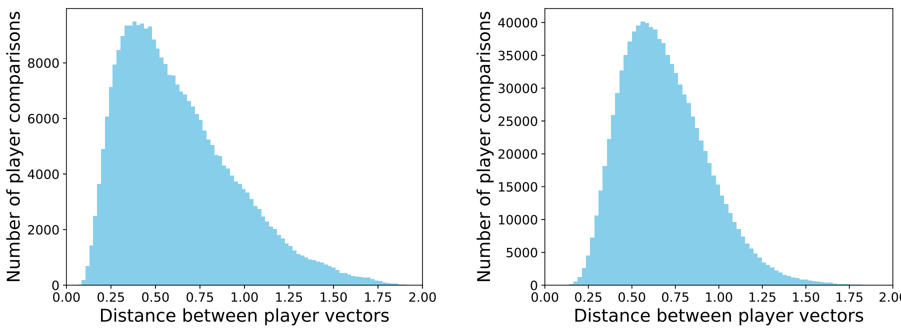
Finally, all features are concatenated together for each player, resulting in player vectors with 13 skill features for defenders and 18 skill features for forwards.

Figs. 2a and 2b display the distribution of the Euclidean distances between player vectors. As discussed before, defender vectors have length 13, while vectors for forwards have length 18. Each defender is compared to each other defender, and similarly for forwards. The distances between these vectors range from approximately 0.1 to 1.75 for both forwards and defenders. Most defenders fall within the range of 0.4 to 0.75, whereas forwards are typically found between 0.4 and 1.0.

5 Playing Style Classification

Given the skill vectors for players, our aim is to generate different categories of playing styles. As we wanted to model that players can take on different

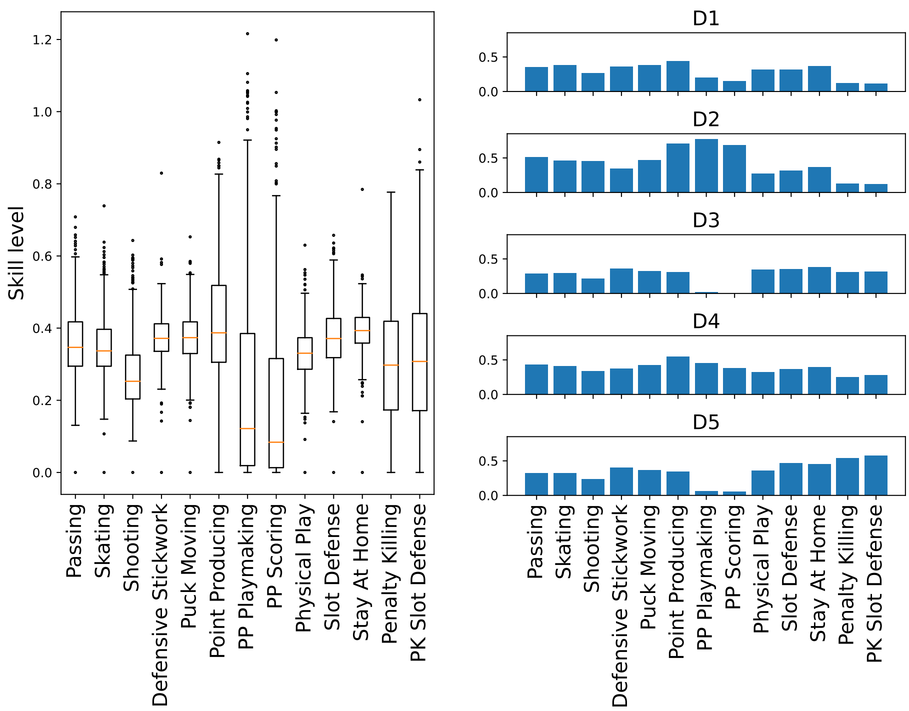
[3] The lower edge of the box represents the lower quartile value (25%) value, the (yellow) line in the box the median (50%) value, and the upper edge of the box the higher quartile (75%) value. The lower whisker shows the minimum value and the upper whisker the maximum value. Points below the lower whisker or above the upper whisker are outliers.



(a) Skill vectors for defenders.

(b) Skill vectors for forwards.

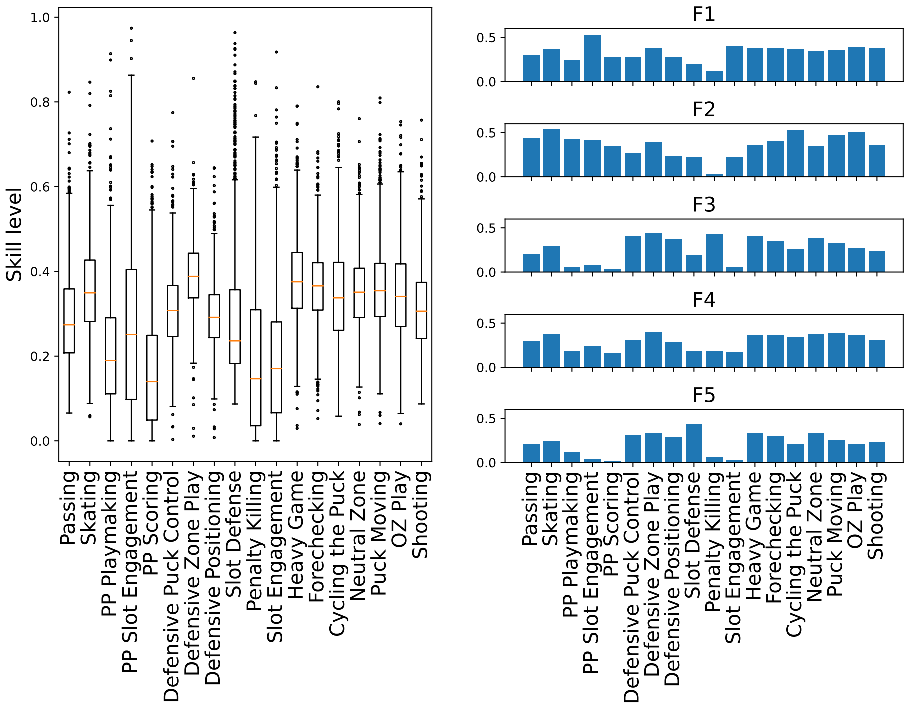
Fig. 2: Distribution of Euclidean distances between player vectors.



(a) Skill distribution (D).

(b) Cluster centroids (D).

Fig. 3: Skill distribution and centroid values (D).



(a) Skill distribution (F). (b) Cluster centroids (F).

Fig. 4: Skill distribution and centroid values (F).

playing styles to certain degrees and that styles may have overlapping elements, we opted to use a fuzzy clustering approach. In this paper, we used the fuzzy c-means algorithm [5,1]. The objective in fuzzy c-means is to create k fuzzy partitions among a set of n objects from a data vector $\mathbf{x}$ by solving (1) until convergence.

$$\min_{\mathbf{U},\mathbf{C}} J_m = \sum_{i=1}^{n}\sum_{j=1}^{k} u_{ij}^m d^2(\mathbf{x}_i, \mathbf{c}_j) \quad s.t. \quad u_{ij} \in [0,1], \sum_{j=1}^{k} u_{ij} = 1 \tag{1}$$

In (1), d denotes the distance between object i and the j:th cluster centroid c_j. Further, u_{ij} is the degree of membership for object i to cluster j. The hyperparameter m controls the degree of fuzziness, where a higher m leads to a fuzzier solution. The fuzzy solution converges to the crisp solution when $m \to 1$. When $m \to \infty$ then $u_{ij} \to \frac{1}{k}$.

For the implementation of fuzzy c-means clustering the open source fuzzy-c-means package for Python was used [4]. We used the maximum possible number (1,000) of iterations. The hyperparameter m was set to 1.5, which was determined by calculating the Fuzzy Partition Coefficient (FPC) that indicates how well the model can divide the data points into clean clusters across different m-values. We investigated which value for k to use with different methods and decided to use 5.

The clustering resulted in five clusters for defenders and five clusters for forwards, where players have certain degrees of membership for each of the clusters of their role. In Tables 3 and 4 we show for each cluster the ten defenders and forwards, respectively, that have the highest degree of membership for that cluster. All of these degrees of membership for the top ten players are over 0.8.

Table 3: Clusters for defenders.

Cluster D1 (91 players)	Cluster D2 (229 players)	Cluster D3 (188 players)	Cluster D4 (128 players)	Cluster D5 (142 players)
S Forsmark (SHL)	R Murphy (AHL)	J Nyberg (SHL)	D Brickley (SHL/HA)	B Pachal (AHL)
H Skinner (AHL)	L Cormier (AHL)	A Söderberg (HA)	F Kral (AHL)	A Strand (AHL)
W Wallinder (SHL/AHL)	T Smith (AHL)	Y Kuznetsov (AHL)	E Sjöström (SHL/HA)	D Samorukov (AHL)
J Andersson (SHL)	L Mailloux (AHL)	K Lowe (SHL)	M Setkov (HA)	I Solovyov (AHL)
H Gabrielsson (HA)	C Carrick (AHL)	P Tischke (AHL)	S Åkerström (HA)	G Brisebois (AHL)
A Brandhammar (HA)	T Niemelä (AHL)	V Pulli (AHL)	M Björk (AHL/SHL)	M Kokkonen (AHL)
H Styf (HA)	A Lindelöf (HA)	J Lundegård (SHL)	K Johansson (HA)	W Aamodt (AHL)
C.J Lerby (SHL/HA)	J Laleggia (SHL)	L Jardeskog (HA)	J Jansson (HA)	M Karow (AHL)
Q Schmiemann (AHL)	A Kniazev (AHL)	H Falk (HA)	J McIsaac (AHL)	D Helleson (AHL)
J Brook (AHL)	J Pudas (SHL)	I Heens (SHL/HA)	O Nilsson (SHL)	S Santini (AHL)

To investigate which skills are important in the different clusters, we used the ten players with highest membership degree from each cluster to compute centroids for the clusters.

In Fig. 3b we show the values for the skills of the centroids in the defender clusters. As these values are based on the skill values of the players with top

Table 4: Clusters for forwards.

Cluster F1 (302 players)	Cluster F2 (359 players)	Cluster F3 (255 players)	Cluster F4 (243 players)	Cluster F5 (250 players)
T Barron (AHL)	L Larsson (SHL)	J Grönhagen (HA)	R Damiani (AHL)	M Strömwall (AHL/SHL)
M Westfält (SHL)	O Sillinger (AHL)	F Nilsson (SHL)	S Walker (AHL)	O Palve (SHL)
N Caamano (AHL)	R Elie (SHL)	F Barklund (HA)	N Todd (AHL)	M Ruohomaa (SHL)
N Jones (AHL)	A Räty (AHL)	J Devane (AHL)	R Marenis (HA)	D Holloway (AHL)
K MacLean (AHL)	J Kellman (SHL)	R Clune (AHL)	A Beckman (AHL)	D Tomasek (SHL)
M Marushev (AHL)	J Lauko (AHL)	R Muzik (SHL)	C Conacher (AHL)	J Looke (SHL/HA)
M O'Leary (AHL)	A Poganski (AHL)	O Pettersson (SHL)	A Andreoff (AHL)	A Petersson (SHL)
B Maxwell (SHL)	G Meireles(AHL)	J Joshua (AHL)	J Doan (AHL)	A Louis (AHL)
T Kaspick(AHL)	P Carlsson (SHL)	K Gabriel (AHL)	B McCartney (AHL)	M Modigs (HA)
J Labate (AHL)	E Desnoyers (AHL)	I McKinnon (AHL)	S Wright (AHL)	L Bristedt (SHL)

membership degrees, they can be seen as the skill levels representing the playing style for a prototypical player for that cluster.[4] Defenders in D1 do not excel in any particular skill, but they also do not rank the worst in any category. The skills with the highest values are those that facilitate point production and puck movement, which suggests a somewhat more offensive than defensive role in the team. D2 defenders are the most offensively skilled defenders that significantly outperform other defenders in point producing and powerplay skills. These defenders show lower values in the defensive skills such as physical play and penalty killing skills. D3 is comparable to D5 in terms of overall defensive capabilities, although D3 defenders demonstrate lower values in all skills than D5 defenders. This suggests that D3 players are more defensive-minded, but not necessarily the top defensive performers. Further, D4 shows high values overall in all skills but excel the most in passing, point producing and powerplay playmaking. This indicates that these defenders can play both in powerplay and boxplay as they excel both in the defense and offense. The strengths for D5 are penalty killing where they have the highest skill level of all playing styles. D1 also has high values in defensive skills such as defensive stickwork, physical play, slot defense and stay at home. D5 shows lower values in more offensive skills such as passing, skating, shooting, puck moving, and point producing. The powerplay skills are almost non-existent indicating that most of these players do not play in powerplay.

In Fig. 4b we show the values for the skills of the centroids in the forwards clusters. The F1 forwards show high values across all skills, with the highest skill level observed in powerplay slot engagement. This indicates that F1 forwards excel in both defensive and offensive skills. F2 forwards are the offensively skilled players with high values in skating, cycling the puck, and offensive zone play, as well as powerplay skills. F3 forwards demonstrate higher values in defensive skills and lower in the offensive skills. F3 also obtains the highest skill level in boxplay out of all playing styles. F4 shows similar skill values as F1 except a bit

[4] The histogram visualization for the values for the centroids for the clusters would show a similar shape if we would have taken the average values of all players in the clusters instead of the average values for the top 10 players in the clusters, although the peaks would be lower. This is also the case for the forward clusters.

lower powerplay skills exchanged for a bit higher boxplay skills. F4 shows decent skill levels in both the offense and the defense. F5 forwards demonstrate lower offensive values together with decent defensive skill levels. F5 has the highest value of all forwards in slot defense indicating a lot of ice time in the defensive zone.

6 Conclusion

In this paper we represented the playing styles for ice hockey defenders and forwards based on skill sets. The skills were computed based on event data. Further, we used fuzzy clustering to define five types of playing styles each for defenders and forwards and showed typical skill levels and example players for these playing styles. This information can be used for scouting and game preparation.

As future work, we will define a new similarity between players based on their membership values to the playing style clusters. This will not only allow for finding players with the same main playing style, but also where the secondary styles are similar. Further, we will investigate in other clustering methods such as Gaussian Mixture Models.

References

1. Bezdek, J.: Pattern Recognition With Fuzzy Objective Function Algorithms. Springer New York, NY (1981). https://doi.org/10.1007/978-1-4757-0450-1
2. Chan, T.C., Cho, J.A., Novati, D.C.: Quantifying the contribution of NHL player types to team performance. Interfaces **42**(2), 131–145 (2012). https://doi.org/10.1287/inte.1110.0612
3. Decroos, T., Davis, J.: Player vectors: Characterizing soccer players' playing style from match event streams. In: Brefeld, U., Fromont, É., Hotho, A., Knobbe, A.J., Maathuis, M.H., Robardet, C. (eds.) Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2019, Würzburg, Germany, September 16-20, 2019, Proceedings, Part III. Lecture Notes in Computer Science, vol. 11908, pp. 569–584 (2019). https://doi.org/10.1007/978-3-030-46133-1_34
4. Dias, M.L.D.: fuzzy-c-means: An implementation of fuzzy c-means clustering algorithm. (May 2019). https://doi.org/10.5281/zenodo.3066222, https://git.io/fuzzy-c-means
5. Dunn, J.: A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. Journal of Cybernetics **3**(3), 32–57 (1973). https://doi.org/10.1080/01969727308546046
6. Franks, A., Miller, A., Bornn, L., Goldsberry, K.: Characterizing the spatial structure of defensive skill in professional basketball. The Annals of Applied Statistics **9**(1) (2015). https://doi.org/10.1214/14-aoas799
7. Gramacy, R.B., Jensen, S.T., Taddy, M.: Estimating player contribution in hockey with regularized logistic regression. Journal of Quantitative Analysis in Sports **9**(1), 97–111 (2013). https://doi.org/10.1515/jqas-2012-0001

8. Gu, W., Foster, K., Shang, J., Wei, L.: A game-predicting expert system using big data and machine learning. Expert Systems with Applications **130**, 293–305 (2019). `https://doi.org/10.1016/j.eswa.2019.04.025`
9. Kaplan, E.H., Mongeon, K., Ryan, J.T.: A Markov Model for Hockey: Manpower Differential and Win Probability Added. INFOR Information Systems and Operational Research **52**(2), 39–50 (2014). `https://doi.org/10.3138/infor.52.2.39`
10. Lambrix, P., Carlsson, N.: Performance metrics for ice hockey accounting for goal importance. In: Lambrix, P., Carlsson, N., Vernblom, M. (eds.) Proceedings of the Linköping Hockey Analytics Conference LINHAC 2022 Research Track. Linköping Electronic Conference Proceedings, vol. 191, pp. 11–25 (2022). `https://doi.org/10.3384/ecp191003`
11. Lehmus Persson, T., Kozlica, H., Carlsson, N., Lambrix, P.: Prediction of tiers in the ranking of ice hockey players. In: Brefeld, U., Davis, J., Van Haaren, J., Zimmermann, A. (eds.) Machine Learning and Data Mining for Sports Analytics. MLSA 2020. Communications in Computer and Information Science, vol. 1324, pp. 89–100 (2020). `https://doi.org/10.1007/978-3-030-64912-8_8`
12. Lindström, P., Jacobsson, L., Carlsson, N., Lambrix, P.: Predicting player trajectories in shot situations in soccer. In: Brefeld, U., Davis, J., Van Haaren, J., Zimmermann, A. (eds.) Machine Learning and Data Mining for Sports Analytics. MLSA 2020. Communications in Computer and Information Science, vol. 1324, pp. 62–75 (2020). `https://doi.org/10.1007/978-3-030-64912-8_6`
13. Liu, G., Schulte, O.: Deep reinforcement learning in ice hockey for context-aware player evaluation. In: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. pp. 3442–3448 (2018). `https://doi.org/10.24963/ijcai.2018/478`
14. Liu, Y., Schulte, O., Li, C.: Model Trees for Identifying Exceptional Players in the NHL and NBA Drafts. In: Brefeld, U., Davis, J., Van Haaren, J., Zimmermann, A. (eds.) Machine Learning and Data Mining for Sports Analytics. MLSA 2018. Lecture Notes in Computer Science, vol. 11330, pp. 93–105 (2019). `https://doi.org/10.1007/978-3-030-17274-9_8`
15. Ljung, D., Carlsson, N., Lambrix, P.: Player pairs valuation in ice hockey. In: Brefeld, U., Davis, J., Van Haaren, J., Zimmermann, A. (eds.) Machine Learning and Data Mining for Sports Analytics. MLSA 2018. Lecture Notes in Computer Science, vol. 11330, pp. 82–92 (2019). `https://doi.org/10.1007/978-3-030-17274-9_7`
16. Macdonald, B.: A Regression-Based Adjusted Plus-Minus Statistic for NHL Players. Journal of Quantitative Analysis in Sports **7**(3) (2011). `https://doi.org/10.2202/1559-0410.1284`
17. Macdonald, B.: An Improved Adjusted Plus-Minus Statistic for NHL Players. In: MIT Sloan Sports Analytics Conference (2011)
18. Nandakumar, N., Jensen, S.T.: Historical perspectives and current directions in hockey analytics. Annual Review of Statistics and Its Application **6**(1), 19–36 (2019)
19. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research **12**, 2825–2830 (2011)
20. Pettigrew, S.: Assessing the offensive productivity of NHL players using in-game win probabilities. In: MIT Sloan Sports Analytics Conference (2015)

21. Routley, K., Schulte, O.: A Markov Game Model for Valuing Player Actions in Ice Hockey. In: Meila, M., Heskes, T. (eds.) Uncertainty in Artificial Intelligence. pp. 782–791 (2015)
22. Säfvenberg, R., Carlsson, N., Lambrix, P.: Simple and practical goal importance metrics for ice hockey. In: Brecht, T., Carlsson, N., Vernblom, M., Lambrix, P. (eds.) Proceedings of the Linköping Hockey Analytics Conference LINHAC 2023 Research Track. Linköping Electronic Conference Proceedings, vol. 201, pp. 40–52 (2023). https://doi.org/10.3384/ecp201.4
23. Säfvenberg, R., Carlsson, N., Lambrix, P.: Identifying player roles in ice hockey. In: Brefeld, U., Davis, J., Van Haaren, J., Zimmermann, A. (eds.) Machine Learning and Data Mining for Sports Analytics. MLSA 2023. Communications in Computer and Information Science, vol. 2035, pp. 131–143 (2024). https://doi.org/10.1007/978-3-031-53833-9_11
24. Sans Fuentes, C., Carlsson, N., Lambrix, P.: Player impact measures for scoring in ice hockey. In: MathSport International 2019 Conference. pp. 307–317 (2019)
25. Schuckers, M.: Draft by Numbers: Using Data and Analytics to Improve National Hockey League (NHL) Player Selection. In: MIT Sloan Sports Analytics Conference (2016)
26. Schuckers, M., Curro, J.: Total Hockey Rating (THoR): A comprehensive statistical rating of National Hockey League forwards and defensemen based upon all on-ice events. In: MIT Sloan Sports Analytics Conference (2013)
27. Schulte, O., Khademi, M., Gholami, S., Zhao, Z., Javan, M., Desaulniers, P.: A Markov Game model for valuing actions, locations, and team performance in ice hockey. Data Mining and Knowledge Discovery **31**(6), 1735–1757 (2017). https://doi.org/10.1007/s10618-017-0496-z
28. Schulte, O., Zhao, Z., Javan, M., Desaulniers, P.: Apples-to-apples: Clustering and Ranking NHL Players Using Location Information and Scoring Impact. In: MIT Sloan Sports Analytics Conference (2017)
29. Thomas, A., Ventura, S.L., Jensen, S., Ma, S.: Competing Process Hazard Function Models for Player Ratings in Ice Hockey. The Annals of Applied Statistics **7**(3), 1497–1524 (2013). https://doi.org/10.1214/13-AOAS646
30. Vincent, C.B., Eastman, B.: Defining the style of play in the nhl: An application of cluster analysis. Journal of Quantitative Analysis in Sports **5**(1) (2009). https://doi.org/10.2202/1559-0410.1133

Puck Possessions and Team Success in the NHL

Miles Pitassi, Tim Brecht, and Mingyue Xie

Cheriton School of Computer Science, University of Waterloo

Abstract. This paper investigates the relationship between puck possession and team success in the NHL, focusing on the games played during the 2023-2024 regular season (up to the All-Star break). The analysis first reveals a moderate correlation (r = 0.56) between average team possession percentage and Average Goal Differential (Avg. GoalDiff). Next, we introduce Average Offensive Zone Possession Time Differential (Avg. OZPTD) as a key metric, defined as the difference between a team's offensive zone possession time and that of their opponents. We find a strong correlation (r = 0.77) between Avg. OZPTD and Avg. GoalDiff, thereby highlighting its relevance in assessing team performance. Our analysis confirms OZPTD's stability, discriminatory power, and independence from existing metrics like Shot Attempt Percentage (SAT%), also known as Corsi. Additionally, we detail a comprehensive methodology for processing and cleaning possession data sourced from the NHL. This methodology underpins our findings and facilitates future research involving player and team possession data.

1 Introduction

In the National Hockey League (NHL), where the margin for victory is small, teams are in constant pursuit of advantages to enhance their chance of success (i.e., winning). The unpredictable nature of hockey compounds the difficulty of identifying and quantifying metrics that genuinely influence outcomes. In this paper we analyze puck possession and its use as a potential indicator of success. This work is further motivated by the premise shared across sports that possessing the ball or puck for significantly more time than the opponent increases your chance of winning.

Research in other sports present mixed results regarding the correlation between possession and team success; some studies affirm a strong correlation [19] [16] [13] [5] [3], while others find no significant relationship [4] [6] [7] [9] [12] [10]. In the NHL, prior investigations into puck possession have mainly relied on manual tracking [17] or metrics approximating possession such as Shot Attempt Percentage (SAT%), also known as Corsi [20]. SAT% (Corsi) measures a team's share of the total shot attempts in a game. The rationale behind SAT% (Corsi) is that a higher number of shot attempts, which can only be credited during a possession, indicates superior puck control.

Despite these indirect methods for measuring possession, the strategic importance of puck possession in the NHL has notably increased, particularly during overtime since the introduction of 3-on-3 overtime in 2015. This growing emphasis is underscored by the NHL general managers convening this season to discuss potential regulatory changes, such as implementing a shot clock during overtime [22]. This consideration directly reflects concerns about extended possessions during overtime, highlighting the central role puck control has come to play in modern NHL strategies.

With the recent introduction of puck and player tracking (PPT) technologies in the 2021-22 NHL season and the significance of puck possession, this paper investigates whether teams with more puck possession have greater success. The contributions of this paper are:

– We provide a methodology for cleaning and processing the NHL's player possession data. This is required to support our analyses and also lays the groundwork for future studies.
– We examine correlations between several possession metrics and indicators of team success, providing insights into their predictive value.
– We introduce the Average Offensive Zone Possession Time Differential (Avg. OZPTD) metric, defined as the difference between a team's offensive zone possession time and that of their opponents. Avg. OZPTD is highly correlated ($r = 0.77$) with Average Goal Differential. This highlights its potential for enhancing our understanding of team success in the NHL.
– We show that Avg. OZPTD's is stable across two halves of our dataset (i.e., useful for prediction), is able to differentiate between teams and is independent from existing metrics which demonstrate its potential as a useful new metric.

2 Related Work

Previous studies on the importance of possession across sports provide context for this paper, particularly highlighting the research gap in hockey analytics.

In football (soccer), the relationship between ball possession and team success has been extensively studied using various research methods. Some of these studies find a positive relationship. For instance, researchers studied the 2016 UEFA Euro and found that the average possession time for a leading team was 20.3 minutes with a standard deviation (SD) of 16.0 minutes, compared to 18.2 minutes with a SD of 16.8 minutes for teams when the score was tied, and 13.7 minutes with a SD of 12.3 minutes for a trailing team [5]. The authors explained that the p-value, which assesses the likelihood that these differences occurred by chance, was less than 0.01, indicating a statistically significant difference. Additionally, researchers studying the 2006 FIFA World Cup found that the percentage of ball possession, analyzed using principal component analysis, had the greatest influence on match outcomes with a coefficient with an absolute value of 0.72. This indicates that it is an important variable for discriminating winning teams from those that lose or draw [19]. Another study found that ball possession had a positive effect on winning in the 2014 FIFA World Cup, with an 11% increase in the probability of winning for all matches and a 14% increase for close matches when ball possession increased by two standard deviations [16]. Also, a study covering the 2017-18 and 2018-19 season in the German Bundesliga showed a positive correlation ($r = 0.75$) between team possession and overall points earned [3].

However, other studies have found that possession may not correlate with or may even negatively impact team success. For example, researchers for the FIFA Training centre studied the 2022 FIFA World Cup and found that, for the men's tournament, teams with less possession than their opponents won slightly more games (26 wins versus 23) [10]. Additionally, a study of the 2010-11 season in the Portuguese Premier

League found that the amount of ball possession had a very weak negative correlation (r = -0.192) with the match result [12]. In a study analyzing elite leagues in Europe, researchers found that a significant difference in ball possession percentages between winning and losing teams only occurred in matches with wide result margins (3 or more goals). In the other, closely contested matches, the difference in possession between winning teams (51.48% with a SD of 13.05%) and losing teams (48.52% with a SD of 13.05%) was not statistically significant [9]. Similarly, researchers studying the World Cups of 2002, 2006 and 2010 found that ball possession was slightly higher for winning teams (51.6% with a SD of 6.8%) compared to those that drew (49.9% with a SD of 5.8%) or lost (48.5% with a SD of 6.8%), though the differences were not statistically significant [6]. In another study using data from five European leagues, UEFA, and FIFA tournaments, researchers found that possession time was a poor predictor of team success once team quality and home advantage were accounted for [7].

In basketball, intuition might lead one to believe that possession is less important due to the shot clock, which mandates a field goal attempt within 24 seconds in the NBA and most European leagues. Previous studies have shown a positive but insignificant correlation between longer possessions and success. Research on the Spanish Basketball Playoffs from the 2004-05 season investigated the possession durations of winning and losing teams against various defensive systems. They found that, when averaged across all defenses, winning teams had an average possession duration of 13.1 seconds with a SD of 6 seconds, compared to 12.32 seconds with a SD of 5.88 seconds for losing teams [13]. Although significant differences were observed depending on the defensive system faced, these differences did not translate into statistically significant overall differences in possession durations between winning and losing teams.

In American football, significant value is placed on time of possession, notably because it allows the defense to rest, enhancing both offensive and defensive performance. Time of possession refers to the amount of game time an NFL offense has the ball. Researchers studied the 2003-04, 2004-05, and 2005-06 NFL seasons and found that 67% of teams with greater time of possession than their opponents won their games [4]. However, the research recognized potential biases; leading teams often prolong their possessions near the end of the game to conserve their lead. To avoid this bias, the analysis was confined to first-half data. In this analysis, a logisitic regression model was applied to predict the halftime score. The model revealed a negative coefficient for time of possession (β = -0.126), indicating that for each additional minute of possession in the first half, the log-odds of winning at halftime decrease by 0.126. This indicates that more possession, with biases removed, does not contribute positively to winning.

Hockey's analysis of puck possession has comparatively been less robust as it relies on manual tracking [17] or metrics approximating possession such as SAT% (Corsi) [20]. Some studies using manual tracking or SAT% (Corsi) have found a positive correlation between possession and team success. For instance, a study of 243 NHL overtime periods from 2015 to 2021 in which possessions were manually tracked revealed that victorious teams in 3-on-3 overtime generally have a higher count of individual possessions (53 percent of the total number of individual possessions of both teams), a higher duration of individual possession (54 percent of the total duration of individual possession of both teams), and more offensive zone time (57 percent of the total offensive zone time of both teams) compared to teams that lost [17]. Additionally, a study of the 2007-08, 2008-09, and 2009-10 NHL regular seasons revealed that SAT% (Corsi) Tied (even strength SAT% (Corsi) with the score tied) is more predictive of how well a team will perform ($r = 0.47$) than goal ratio ($r = 0.35$) or winning percentage ($r = 0.34$) [20]. This correlation is relatively low compared to our findings, where higher correlations emerge from utilizing PPT data to measure various metrics of puck possession, most notably for Average Offensive Zone Possession Time Differential (Avg. OZPTD).

Although previous hockey analytics research shows a positive correlation between possession and team success, there are challenges to manually tracking possession. As well, SAT% (Corsi) has its limitations, as it does not account for possession in the defensive or neutral zones and may not reflect the strategy of teams that prioritize high-quality shots over quantity.

In recent years, expected goal (xG) models have gained popularity. Originating in football (soccer), xG represents the probability that a scoring opportunity will result in a goal. It addresses some issues with SAT% (Corsi) as it includes weighting shot attempts based on quality, recognizing that certain shots have a higher probability of resulting in a goal. In hockey, efforts to evaluate shot quality began in 2004 [23] [14] [15]. This foundational work led to the first explicit mention of xG in hockey in a 2012 study, which used ordinary least squares (OLS) regression and ridge regression to predict goals, incorporating variables such as goals, shots, missed shots, blocked shots, faceoffs, hits, turnovers, and zone starts [18]. Since 2012, numerous xG models have emerged, each aiming to capture the best set of predictive variables, often including more than ten variables weighted during model training [8] [24] [26] [25]. These models generally outperform SAT% (Corsi) and other metrics in predictive accuracy [8] [25].

However, there are drawbacks to xG models. First, there are many different xG models, which can have varying parameters, potentially leading to inconsistencies when advising a team on how to improve their xG to win more games. Additionally, to our knowledge, there has been limited work on testing the stability of these models, meaning the parameters and weights might not remain consistent from one season to the next. Lastly, because these models have several parameters, determining the specific actions a team can take to improve their xG may not be straightforward.

In this paper, we utilize PPT data to conduct a detailed investigation into measures of puck possession and their correlation with NHL team success. Our findings indicate that a single metric of possession can be as effective, if not more so, than existing, publicly available xG models in predicting team success.

3 Background

3.1 Definitions of Individual and Team Puck Possession

Before we delve into the dataset and analysis, we define the concepts of individual and team puck possessions as utilized in our study.

According to the NHL definition for the model that produces the individual possession data we employ, a player is considered to have possession and control of the puck, and thus in individual possession, when they make two or more consecutive touches with the puck. The start of the individual possession is marked by the first touch, which is confirmed upon a second touch. Individual possessions also includes brief moments during one-touch actions, like shots, passes, or area plays (e.g., dump-ins). An individual possession ends when the player is separated from the puck or when another player gains possession. We delineate these episodes to identify windows of time with "no individual possession", representing segments of active gameplay where the puck is not under direct control by any player. This includes scenarios ranging from face-offs, puck battles, and loose pucks to passes, shots, and "area plays" (e.g., dump-ins and dump-outs). The top line in Figure 1 shows examples of individual possessions by members of different teams (red and blue lines) and "no possession" (orange dotted lines).

We define team puck possession as the aggregate of individual possessions with continuous possession by members of the same team, interrupted only by game stoppages or a change in possession to the opposing team. Consequently, "no team possession" intervals are distinct from "no individual possession" intervals. The bottom line in Figure 1 shows examples of team possession (red and blue lines) and "no team possession" (green dashed lines). As shown in the figure, team possessions end when the puck is last touched by one team, prior to the opposing team gaining possession. Our use of team possession differs slightly from the official NHL definition as the details required to implement the NHL's definition aren't available in the current dataset.

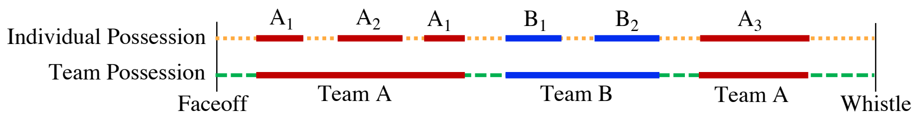


Fig. 1. Differentiating individual and team possessions

3.2 Dataset Overview

This paper utilizes the proprietary PPT dataset provided to us by the NHL. The PPT data, captured through devices in players' sweaters and the puck, records x, y, and z coordinates at high frequencies: 60 times per second for the puck and 12 times per second for each player on the ice. An additional update is provided once per second for players on the bench, resulting in around 734,400 location points in a typical 60-minute game. Due to the varied frequencies of data collection across players and the puck, and the lack of synchronization between devices, all player and puck positions are interpolated to uniform timestamps every hundredth of a second.

In March 2023, a significant advancement was made with the introduction of an individual player possession model into the "DISH" data stream, which features Delayed, Interpolated, Smoothed and Hundred-Hertz enhancements. This dataset is considered unofficial by the NHL and may differ from other datasets that track possession information (e.g., a hand-labeled dataset). Our study uses the DISH data to compute team possessions, which form the basis of our analysis. Since this data only became available in March 2023, the dataset for the 2022-2023 season is limited. Consequently, our analysis focuses on the 2023-2024 NHL season, using data from 780 games played up to January 31, 2024 (the All-Star break). After excluding games with significant data issues, as detailed in Section 4.3, or those with no tracking data, such as the Heritage Classic and the games played in Europe, 708 games remained for analysis.

4 Dataset Cleaning and Filtering

The player possession data provided in the NHL's DISH data stream, indicating who held the puck and for how long, lacks broader game context such as powerplay situations and player locations. To address this, we merge it with data from a detailed game information file, enriching player possessions with relevant game context, and then compute team possessions based on this integrated data. Through this process, we encounter challenges that necessitate extensive cleaning and preprocessing to ensure data integrity. Cleaning refers to the process of correcting or removing inaccuracies within the data that can be rectified, such as adjusting timestamps or eliminating duplicates. Conversely, filtering is our strategy for dealing with more complex issues that cannot be directly corrected; it involves the exclusion of entire games from our dataset.

4.1 Dataset Cleaning

Possessions Occurring During Stoppages: One issue with the data is that some possessions occur either entirely during stoppages or span active and stopped intervals. These erroneous possessions are identified after merging the player possession file with the game information file and computing active gameplay intervals. To resolve this issue, we eliminate portions of the possession that occurred during stoppages, ensuring accuracy in active play representation.

Abnormal Timestamps and Non-chronological Data Entries: The game information file contains updates every hundredth of a second, but some of these updates display additional digits of precision and are out of sequence. These extraneous updates, found to be non-essential, are removed to maintain dataset integrity. After their removal, the data is re-sequenced to reflect the actual gameplay order.

Clock Resets: Another issue encountered in the dataset are unexpected time jumps, with the time remaining on the scoreboard clock suddenly increasing, leading to duplicated timestamps. These time jumps primarily occur after video reviews where time is added back to the clock, such as when a play is subsequently ruled as offside. Smaller adjustments may also result from false face-offs or if the clock inadvertantly continues running briefly after a whistle. The NHL addresses these situations by eliminating all recorded statistics and events that transpired during the time that is later nullified. Our approach mirrors this; upon identification of such a clock reset, we disregard stats and possessions recorded during the time frame subject to the reset.

Overlapping and Duplicate Player Possessions: The last challenge rectified through cleaning is the presence of duplicate or overlapping possessions. Duplicates are resolved by retaining a single entry. For overlapping possessions, we evenly distribute the overlapping time (i.e., the period during which the data indicates two players simultaneously possess the puck) among the involved players.

4.2 Dataset Filtering

There are cases where the above cleaning methods are insufficient to repair the data and preserve the integrity of the dataset. Consequently, we establish exclusion criteria based on the severity of data corruption: if the data is compromised for either more than 4% of a game's duration or more than 4% of a team's possession time, we exclude the game from our analysis. This filtering process results in the exclusion of 68 games, leaving 91% of the games for which we have data available for use in our analysis. The distribution of team appearances in the excluded games varied, with an average of 4.5 games per team, a standard deviation of 2.2 games, a minimum of 1 game for the Tampa Bay Lightning (TBL), and a maximum of 11 games for the Vancouver Canucks (VAN), constituting 22.4% of their total games. In Section 6.1, we show that robust analysis can be achieved with just 20% of a team's games in our dataset, as the correlation between early game data and the rest of the season stabilizes after this. Table 1 shows the number and percentage of games impacted by each filter; note that the sum of games exceeds 67 and the sum of percentages exceeds 9% since 20 games were subject to more than one filter criterion.

Filter Criteria	**Games Impacted**	**Percent Impacted**
Irregular Possession Lengths	34	4.4%
Clock Gaps	30	3.8%
Irregular Period Lengths	26	3.3%
Possessions with Missing Data	5	0.6%
Excessive Distance Between Puck and Possessor	2	0.3%

Table 1. Impact of various filters on game dataset

Irregular Possession Lengths: Games are flagged for exclusion when the duration without any possession or the length of specific possessions significantly exceeds normal expectations. For total "no possession" time, we apply the statistical outlier definition of mean plus three standard deviations. Given the mean of 62.8% and the standard deviation of 4.8%, this led to the exclusion of any game exceeding 77.2%. Additionally, games with a no possession duration longer than 144 seconds, or any individual possession lasting more than 48 seconds, are excluded, impacting 4.4% of the total games. The limit of 144 seconds corresponds to 4% of a 60-minute game and 48 seconds represents 4% of the average of the per game sums of individual possession times (20 minutes).

Clock Gaps: We identify games with significant windows of time missing in scoreboard data timestamps, indicating lost data segments affecting puck locations, player locations, or possession details. We set a 144-second threshold for these gaps, equivalent to 4% of a 60-minute game. Games exceeding this limit due to missing data are excluded, affecting 3.8% of the dataset.

Irregular Period Lengths: We identify games with periods deviating significantly from the standard 20-minute length in order to filter games with extensive data loss or situations where our data cleaning techniques may be ineffective. We exclude games with periods exceeding or falling short of the expected duration by more than 48 seconds, equivalent to 4% of a 20-minute period, impacting 3.3% of the total games.

Possessions with Missing Data: Games are flagged when they contain missing player data, or missing possession start or end times. This is likely due to tracking failures in the puck or jerseys, or instances where a player does not have a tracking device in their jersey. Games with more than two instances of missing data related to possessions are removed from the dataset; impacting 0.6% of our games.

Excessive Distance Between the Puck and Possessor: We considered possessions where the distance between the puck and its possessor is too large. We focus on possessions where the puck is over 16 feet from the possessor continuously for more than 2 seconds, indicating potential data inaccuracies. Games with a total "excessive distance duration" exceeding 48 seconds, equivalent to 4% of a team's average possession time of 20 minutes, are excluded, affecting 0.3% of the total games. In previous work, we adjusted the timestamps for events like shots and passes to try to more accurately capture the point of release [21]. We considered a similar approach in this work but the problem proved more difficult because we found instances where the distance between the puck and possessor is large in the middle of the possession. Adjusting such possessions would amount to building a new model, which is currently the domain of the NHL.

5 Analysis of Team Possessions

In this paper, we explore the relationship between team success and possession metrics, focusing on team possession percentage, aggregate individiual possession count differential, and offensive zone team possession time differential. Team success is measured primarily by goal differential because it is adaptable across game situations, unlike points per game, which is less flexible. Additionally, for the games in our dataset, average goal differential exhibits a strong correlation with average points per game (r = 0.95). Note that, unless stated otherwise, the analysis includes all strengths (i.e., even-strength and powerplays) and pertains exclusively to regulation time. This means that for our analysis, each team is awarded one point if a game goes to overtime.

5.1 Team Possession Percentage Versus Team Success

Team possession percentage is calculated by dividing the total duration of team A's possession by the combined possession duration of team A and the opposing team. Team possession percentage is calculated for each team in every game and subsequently averaged across all games played. We compute the correlation between average team possession percentage and average goal differential (Avg. GoalDiff), as well as average goals for (Avg. GF) and average goals against (Avg. GA), aiming to delineate the correlations of possession with offensive and defensive metrics.

As shown in Table 2, average team possession percentage is moderately correlated with Avg. GF, suggesting that teams with higher possession tend to score more goals. In contrast, the correlation between average team possession percentage and Avg. GA is weaker, implying that while possession might play a role in limiting opposition goals, its effect is not as strong.

Possession Metric	**Success Metric**	***r*-value**
Avg. Team Possession Percentage	Avg. GF	0.56
Avg. Team Possession Percentage	Avg. GA	-0.38
Avg. Team Possession Percentage	Avg. GoalDiff	0.56

Table 2. Correlations between average team possession percentage and team success metrics

Furthermore, our analysis reveals a nonlinear relationship among the correlations of average team possession percentage with Avg. GF, Avg. GA, and Avg. GoalDiff. Intuition might lead one to expect these correlations to sum linearly; for example, given the correlation between average team possession percentage and Avg. GF is +0.56, and between average team possession percentage and Avg. GA is -0.38, one might anticipate the correlation between average team possession percentage and Avg. GoalDiff to be the difference, equating to +0.94. This is not true and can be explained by understanding the correlation formula's normalization process. The Pearson correlation coefficient is:

$$r = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \tag{1}$$

where $\text{cov}(X, Y)$ is the covariance between X and Y, and σ_X and σ_Y are the standard deviations of X and Y, respectively. The denominator normalizes the covariance by dividing it by the product of the standard deviations of X and Y, ensuring the correlation values fall within the range of -1 to +1. Given the distinct standard deviations for Avg. GF (0.44), Avg. GA (0.39), and Avg. GoalDiff (0.70), this normalization introduces nonlinearity to the relationships.

5.2 Possession Count Differential Versus Team Success

Shifting our analysis from the percentage of team possession to the aggregate quantity of individual possession instances can potentially offer new insights by capturing both the totality of possessions gained through turnovers or puck battles and the extent of puck movement within team possessions. To assess which teams excel in managing aggregate individual possession quantity, we introduce a metric called average possession count differential.

For team A, the possession count differential is defined as the count of team A's individual possessions, minus the count of the opposing team's individual possessions. We compute this metric for each game and subsequently determine the average across all games played by each team. Utilizing this metric reveals a slightly enhanced correlation with Avg. GoalDiff (r = 0.63) compared to the correlation between average team possession percentage and Avg. GoalDiff (r = 0.56). This improved correlation may indicate the potential impact of frequent and dynamic possession changes to outscoring opponents, suggesting a strategy centered on maximizing possession instances correlates positively with achieving a better goal differential.

5.3 Offensive Zone Possession Time Differential Versus Team Success

We now examine the significance of possession within the offensive zone. The rationale for this approach is that possessions in the defensive or neutral zones can serve to facilitate transitions, whereas offensive zone possessions might contribute more directly to scoring goals and outscoring the opponent. In this refined analysis, we introduce a new metric, Offensive Zone Possession Time Differential (OZPTD), which is defined as the sum of the duration of team A's individual possessions in the offensive zone, minus the sum of the duration of the opposing team's individual possessions in their offensive zone (team A's defensive zone). For possessions that span multiple zones, the duration is allocated proportionally based on the time spent in each zone. Similar to the previously examined metrics, OZPTD is computed for each game and subsequently averaged across all games played by each team.

As shown in Figure 2, our analysis reveals a significant positive correlation of 0.77 between Avg. OZPTD and Avg. GoalDiff. This finding highlights the importance of not just possessing the puck more than the opponent, but doing so in the offensive zone where it more strongly correlates with outscoring the opponent. Teams such as the Colorado Avalanche (COL) and Florida Panthers (FLA) who, on average, maintain offensive zone possession longer than their opponents, typically see positive goal differentials. Interestingly, the Winnipeg Jets (WPG), Boston Bruins (BOS) and Vancouver Canucks (VAN) achieved the highest Avg. GoalDiff values despite having values of Avg. OZPTD near the league average of 0. In contrast, the San Jose Sharks (SJS) and Chicago Blackhawks (CHI) exhibit negative Avg. OZPTD values and, correspondingly, negative Avg. GoalDiff values. Recognizing that SJS and CHI may contribute significantly to the strong correlation, we compute the correlation coefficient without those two teams and observe an r-value of 0.63. In future work we plan to examine if offensive zone possession counts and differential are also correlated with success.

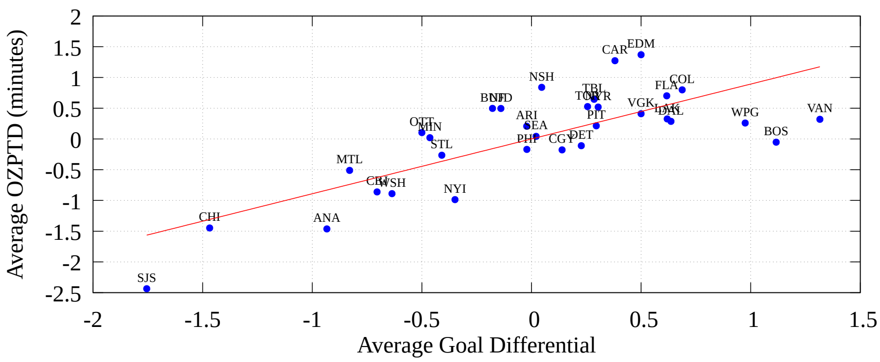


Fig. 2. Avg. OZPTD versus Avg. GoalDiff (r = 0.77)

As Avg. GoalDiff is highly correlated (r = 0.95) with average points per game for the games in our dataset, Figure 2, which arranges teams from left to right based on Avg. GoalDiff, provides a useful reference for readers to assess team standings, offering a more accurate perspective than actual standings that include games outside our analysis.

Recognizing the significance of the correlation, we conduct a deeper examination of its components, focusing exclusively on even-strength play. The correlation remains high at 0.73, indicating that the initial correlation is not simply a byproduct of power plays but is also prevalent during even-strength play, reinforcing the importance of offensive zone control throughout the game.

Given Avg. OZPTD's strong correlation with Avg. GoalDiff, we also analyzed it on a per-game basis, as depicted in Figure 3. This per-game analysis shows an r-value of 0.00, indicating no correlation. This finding suggests that, despite the correlation between Avg. OZPTD and Avg. GoalDiff across many games, individual games show high variability. Thus, while superior offensive zone possession doesn't guarantee game victories, teams with consistently higher offensive zone time may outscore their opponents over the course of a season.

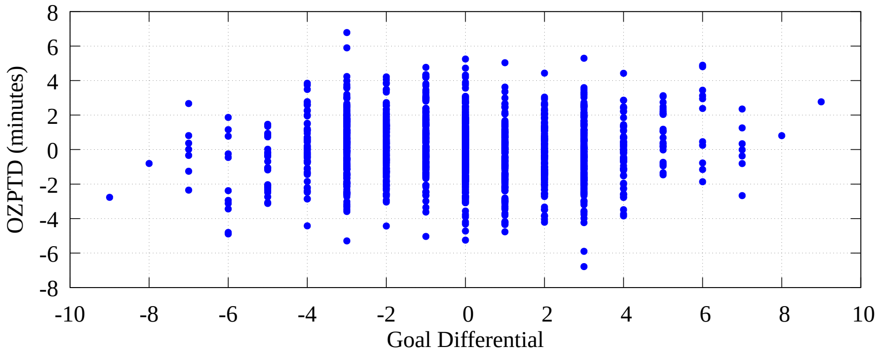


Fig. 3. OZPTD versus GoalDiff for all games (r = 0.00)

5.4 Possession Across Different Strengths

Building on our earlier findings, this section delves deeper into possession metrics across different strength scenarios, as shown in Figure 4. We observe that at even strength, possession is typically balanced between teams. However, with a plus-1 strength advantage, teams dominate possession. In contrast, a minus-1 strength differential leads to a substantial decrease in possession percentage for the disadvantaged team.

The variance in average team possession percentages is notably higher in even strength scenarios than in situations of plus-1 or minus-1. Specifically, the Chicago Blackhawks (CHI) and San Jose Sharks (SJS) show lower possession percentages at even strength, yet they are near the league average in plus-1 and minus-1 situations.

6 Meta Metrics: Evaluating Average OZPTD

Due to Avg. OZPTD's significant correlation with Avg. GoalDiff, and thus its potential to offer insights, it is imperative to evaluate this new metric. We utilize the notions introduced by Franks et al. [11], which emphasizes three key properties: stability, discrimination, and independence. While some of our tests of these properties differ slightly from those suggested in their paper, we maintain the spirit of each property. Stability measures the consistency of a metric across seasons or portions of a season (e.g., the

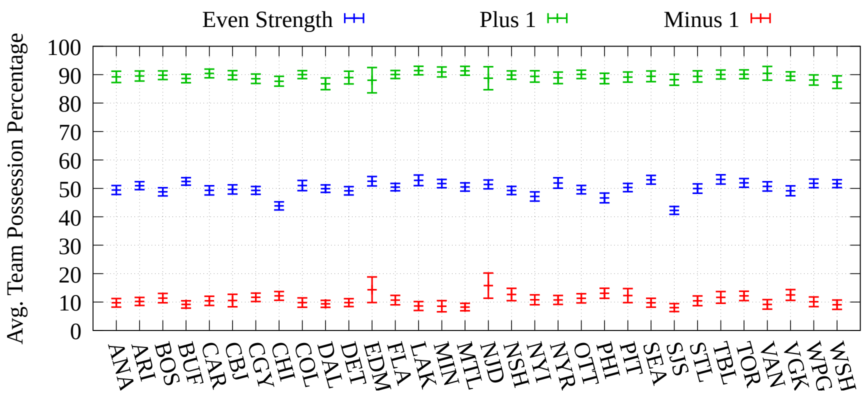


Fig. 4. Average team possession percentage: even strength, +1 and -1 (95% confidence intervals)

value of using the metrics in predictions), discrimination measures its ability to distinguish between players or teams, and independence assesses whether it provides unique insights when compared with existing metrics.

6.1 Stability

To assess the stability of Avg. OZPTD and determine its potential for predictive use, we calculate Avg. OZPTD separately for the first and second halves of the dataset. The observed strong correlation (r = 0.84) between Avg. OZPTD in the two halves, depicted in Figure 5, validates the metric's consistency. To further our understanding of the metric's stability, we conduct a rolling correlation analysis where the Avg. OZPTD is calculated for each team across incremental segments of the dataset, ranging from 5% to 50% and then these values are compared with Avg. OZPTD for the remaining games. The correlation starts at 0.68 when using the first 5% of the games to predict the Avg. OZPTD of the remaining 95% of the games and stabilizes above 0.80 when using the first 20% of the games to predict the remaining 80% of the games.

Predictive Power: To evaluate the predictive accuracy of Avg. OZPTD, we divided our dataset into two halves. Using data from the first half of our dataset, we built a linear regression model to establish the relationship between Avg. OZPTD and Avg. GoalDiff. We then tested this model with data from the second half of our dataset, using measured Avg. OZPTD to predict Avg. GoalDiff for each team. Our predictions were compared to the actual outcomes, resulting in an R^2 value of 0.49, and a correlation coefficient of 0.73. This correlation indicates a relatively strong correlation between the predicted and actual values.

To compare results obtained when using other metrics for predicting team success, we find that the 2012 study by Macdonald yield a correlation between actual and predicted goals of 0.69 using his ridge regression model [18]. Next, we found that a 2022 study, described on the Hockey-Statistics website [25], that reports that using their xG model and the xG model from Evolving-Hockey [1] to predict the expected goals for

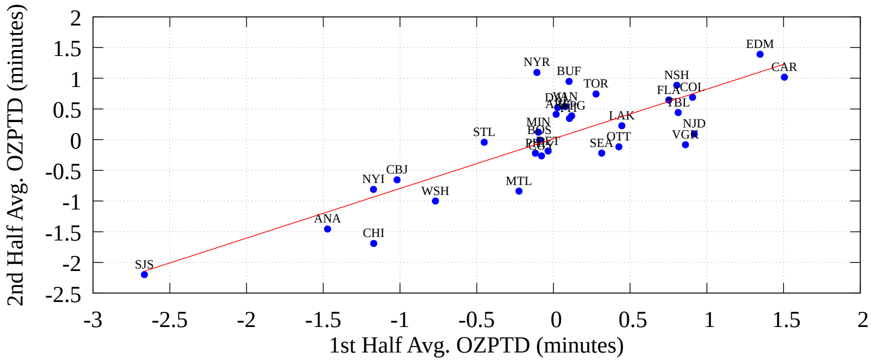


Fig. 5. Average OZPTD across dataset halves (r = 0.84)

percentage (xGF%) yielded an R^2 value of 0.49 in both cases. This prediction was based on the xGF% from the first 41 games of a season for each team to forecast the xGF% for the last 41 games. Note that xGF% is the ratio of a team's expected goals for compared to their opposition. Another study from 2015 considers a different expected goals model, and found that when using their model built using the first 40 games to predict GF for each team at the end of the season, they obtained an R^2 value of 0.51 [8].

While a more in-depth evaluation needs to be done using larger sample sizes with a direct comparison between metrics, this preliminary investigation indicates that our fairly simple Avg. OZPTD metric performs on par with existing, relatively complex models (because they typically use a large number of parameters that appear to require tuning) for predicting team success.

6.2 Discrimination

Our evaluation of Avg. OZPTD's discriminatory power, depicted in Figure 6, shows the Avg. OZPTD for each team, including 95% confidence intervals. There are statistically significant differences between some teams, however the overlap in confidence intervals for many teams indicates that the metric might have moderate discriminatory power.

6.3 Independence

In assessing the independence of Avg. OZPTD, we revisit SAT% (Corsi) and expected goals (xG). SAT% (Corsi) has traditionally been used to approximate possession by measuring the ratio of a team's shot attempts (goals, shots on net, shots that miss the net, and blocked shots) to the total shot attempts in the game. xG models attempt to improve on the predictive power of SAT% (Corsi) by including several variables related to the shot to better describe the context around the shot. To analyze the independence of Avg. OZPTD from these two metrics, we show the correlation between them, as well as each metric's correlation with team success (Avg. GoalDiff) as shown in Table 3. Note that the data used for the xGF% model is from Natural Stat Trick [2].

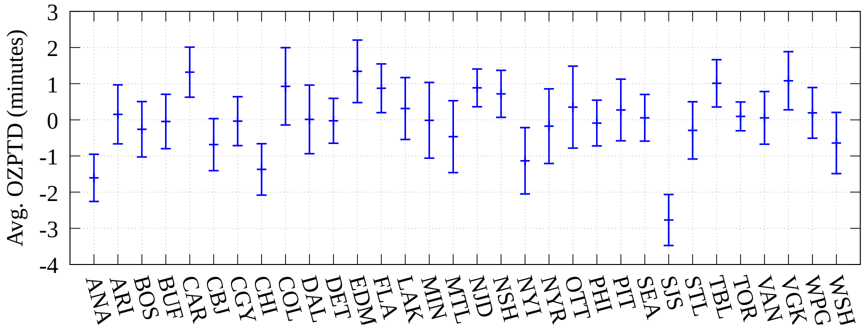


Fig. 6. Average team OZPTD, with 95% confidence intervals

The results indicate that Avg. OZPTD is strongly correlated to SAT% (Corsi) but shows a stronger correlation to Avg. GoalDiff compared to the correlation between SAT% (Corsi) and Avg. GoalDiff. This stronger correlation for Avg. OZPTD implies it provides additional insights beyond SAT% (Corsi), especially in relation to game outcomes. The results also indicate that Avg. OZPTD is strongly correlated to xGF%, with both metrics having the same correlation to Avg. GoalDiff. However, as mentioned previously, there are potential drawbacks to xG models, such as possible inconsistencies in parameters across different models and the complexity of determining specific actions to improve xG.

Metric	**Correlation with Avg. OZPTD**	**Correlation with Avg. GoalDiff**
Avg. OZPTD	1.00	0.77
SAT% (Corsi)	0.83	0.62
xGF%	0.88	0.77

Table 3. Correlation of Metrics with Avg. OZPTD and Avg. GoalDiff

In terms of evaluating established metrics, the work by Franks et al. [11] does evaluate some NHL metrics for individual players but does not include team metrics. As described in Section 6.1, some previous studies have examined the predictive power of various expected goal (xG) models. However, there is a lack of work in evaluating those metrics in terms of stability, discrimination and independence. In the future, we hope to evaluate established team performance metrics alongside our metrics.

7 Conclusions

In this paper we examine team possession metrics and whether they correlate with team success. Interestingly, we find that average team possession percentages are only weakly correlated with team success metrics like, average goals for ($r = 0.56$) and average goal differential (also $r = 0.56$). We introduce a new metric called the average offensive zone possession time differential (Avg. OZPTD) which measures the difference between the time that team A has possession of the puck while in their offensive zone and the time that the opposing team has possession of the puck while in their offensive zone (i.e., team A's defensive zone). We find that there is a strong correlation between Avg. OZPTD and Avg. GoalDiff (r = 0.77). Furthermore, we show Avg. OZPTD to be stable, capable of discriminating between teams, and providing new information over other metrics like SAT% (Corsi). The strong correlation and these attributes underscore its potential to provide deeper insights into team success.

The existence of the NHL's possession data paves the way for more detailed and exciting analysis. With our methodology for preparing, cleaning, and filtering possession data, we are poised to further investigate possessions in future work. On the team-level, it would be interesting to determine if time spent in the offensive zone correlates with team success or if puck possession is a key component. We would also like to examine chains (or sequences) of individual possessions. Metrics of interest would be the length of the chain and the number of different players in the chain. We would also like to study individual player possessions and correlations with player and team success.

Acknowledgments

We thank Ben Resnick from the National Hockey League's Research and Development Team and Jonah Eisen from Rogers Communications for fruitful discussions and feedback related to this work. We also thank the anonymous reviewers and Dave Radke for their constructive and helpful feedback. We thank Rogers Communications for providing funding for this research. In addition, this project is partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) and a University of Waterloo Undergraduate Research Fellowship. We thank Neel Dayal from Rogers Communications and the National Hockey League's Information Technology, and Stats and Information Teams for making this research possible. We also thank AWS for providing us with credits and technical assistance.

References

1. Evolving Hockey, `https://evolving-hockey.com/`
2. Natural Stat Trick, `https://naturalstattrick.com/`
3. Andrez, P., Silva, M., Carling, C.: Analysis of team success based on match technical and running performance in a professional soccer league. BMC Sports Science **14**, 82 (2022)
4. Arkes, J.: Is controlling the rushing or passing game the key to NFL victories? The Sport Journal **14**, 1 (2011)
5. Casal, C.A., Maneiro, R., Ardá, T., Marí, F.J., Losada, J.L.: Possession zone as a performance indicator in football: The game of the best teams. Frontiers in Psychology **8**, 1176 (2017)
6. Castellano, J., Casamichana, D., Lago, C.: The use of match statistics that discriminate between successful and unsuccessful soccer teams. Journal of Human Kinetics **31**, 139–147 (2012)
7. Collet, C.: The possession game? A comparative analysis of ball retention and team success in European and international football, 2007-2010. Journal of Sports Sciences **31**(2), 123–136 (2013)

8. DTMAboutHeart: Expected Goals are a Better Predictor of Future Scoring than Corsi and Goals (2015), https://hockey-graphs.com/2015/10/01/expected-goals-are-a-better-predictor-of-future-scoring-than-corsi-goals
9. Evangelos, B., Aristotelis, G., Ioannis, G., Stergios, K., Foteini, A.: Winners and losers in top level soccer. How do they differ? Journal of Physical Education and Sport **14**(3), 398–405 (2014)
10. FIFA Training Centre: Part 1.1: The relationship between possession and winning (2023), https://www.fifatrainingcentre.com/en/community/fifa-research/possession-analysis/part-1-1-the-relationship-between-possession-and-winning.php
11. Franks, A.M., D'Amour, A., Cervone, D., Bornn, L.: Meta-analytics: Tools for understanding the statistical properties of sports metrics. Journal of Quantitative Analysis in Sports **12**, 151–165 (2016)
12. Gama, J., Dias, G., Couceiro, M., Sousa, T., Vaz, V.: Networks metrics and ball possession in professional football. Complexity **21**(S2), 342–354 (2016)
13. Ángel Gómez, M., Tsamourtzis, E., Lorenzo, A.: Defensive systems in basketball ball possessions. International Journal of Performance Analysis in Sport **6**, 98–107 (2017)
14. Krzywicki, K.: Shot Quality Model: A Logistic Regression Approach to Assessing NHL Shots on Goal (2005)
15. Krzywicki, K.: NHL Shot Quality 2009-10: A Look at Shot Angles and Rebounds (2010)
16. Liu, H., Gomez, M.Á., Lago-Peñas, C., Sampaio, J.: Match statistics related to winning in the group stage of 2014 brazil fifa world cup. Journal of Sports Sciences **33**(12), 1205–1213 (2015)
17. Luszczyszyn, D.: Inside 3-on-3 OT: Strategies and stats for the NHL's most chaotic game state (2021), https://theathletic.com/1672268/2021/02/08/inside-3-on-3-ot-strategies-and-stats-for-the-nhls-most-chaotic-game-state
18. Macdonald, B.: An Expected Goals Model for Evaluating NHL Teams and Players. In: MIT Sloan Sports Analytics Conference (2012)
19. Moura, F.A., Martins, L.E., Cunha, S.A.: Analysis of football game-related statistics using multivariate techniques. Journal of Sports Sciences **32**(20), 1881–1887 (2014)
20. Purdy, T (also known as JLikens): Loose Ends - Part I: Predicting Future Success (2011), https://objectivenhl.blogspot.com/2011/03/loose-ends-part-i-predictive-validity.html
21. Radke, D., Lu, J., Woloschuk, J., Le, T., Radke, D., Liu, C., Brecht, T.: Analyzing passing metrics in ice hockey using puck and player tracking data. In: Proceedings of the Linköping Hockey Analytics Conference. pp. 25–39 (2023)
22. Rosen, D.: OT rule changes, face-off clarity discussed at GM meetings (2023), https://www.nhl.com/news/overtime-rule-changes-face-off-clarity-discussed-at-gm-meetings
23. Ryder, A.: Shot Quality: A Methodology for the Study of the Quality of a Hockey Team's Shots Allowed (2004)
24. Shomer, H.: Analyzing the Game of Hockey: Evaluating my Shooter xG model (2018), https://fooledbygrittiness.blogspot.com/2018/03/evaluating-my-shooter-xg-model.html
25. Skytte, L.: Hockey-Statistics: Building an xG model - v. 1.0 (2022), https://hockey-statistics.com/2022/08/14/building-an-xg-model-v-1-0
26. Younggren, J., Younggren, L.: A New Expected Goals Model for Predicting Goals in the NHL (2021), https://evolving-hockey.com/blog/a-new-expected-goals-model-for-predicting-goals-in-the-nhl

Industry papers

NEXT-LEVEL ICE HOCKEY INSIGHTS WITH KINEXON SPORTS AND DARTFISH

KINEXON Sports and Dartfish announced a significant partnership earlier this year, combining their advanced tracking technologies to enhance analytics in ice hockey. This collaboration aims to deliver new insights and perspectives for coaches, analysts, and fans by integrating their respective player and puck tracking systems. KINEXON Sports, known for its cutting-edge LPS (Local Positioning System) sensor technology, and Dartfish, a leader in AI-driven optical tracking and analytics, are working together to create a new service that promises to transform the landscape of ice hockey analytics. The partnership brings together KINEXON's precision in player tracking with Dartfish's expertise in puck tracking, resulting in a comprehensive solution that offers a deeper understanding of the game.

The integration leverages KINEXON's real-time, highly accurate player tracking with Dartfish's advanced algorithms for puck movement analysis. This combined approach aims to provide teams and analysts with detailed insights into player performance and game dynamics, enhancing their ability to analyze plays, strategies, and outcomes on the ice.

One of the key innovations of this collaboration is the use of sensor fusion technology, which merges Ultra-Wideband (UWB) and optical tracking to deliver a fully automated tracking solution for ice hockey. This development is expected to offer a level of analysis previously unattainable, helping teams to optimize performance and refine tactical approaches.

The service is set to benefit ice hockey teams and leagues by offering a suite of tools designed to enhance player performance, improve strategic planning, and engage fans through enriched content. By providing real-time data and sophisticated analytics, this partnership between KINEXON Sports and Dartfish is poised to redefine the future of ice hockey analytics and entertainment.

More info: www.kinexon-sports.com / www.dartfish.com

Contact:

Reimut Hellmerichs, VP Business Development KINEXON Sports, Reimut.hellmerichs@kinexon.com

Jean Sebastien Merieux, CEO Dartfish, jean-sebastien.merieux@dartfish.com

> “Hudl has been a vital part of our team's success in the NHL for years. The software increases our productivity through the ability to instantly generate reports, all while organizing and labeling high quality video.”

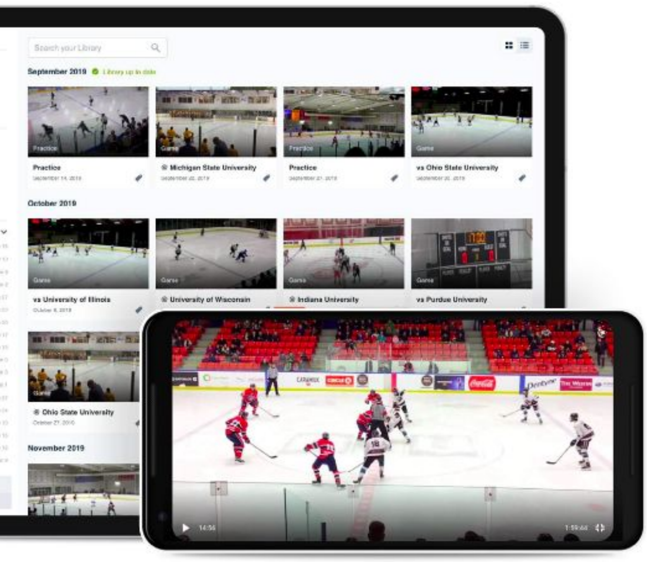
Brett Leonhardt,
Assistant Coach, Washington Capitals

Streamline your workflow

ANALYSIS SOLUTIONS

Analysis For Every Workflow

Online, offline and real-time video and data tools.
- Powerful coding tools
- Customized scripting capabilities
- Video review from any device



DISTRIBUTION SOLUTIONS

Effortless Asset Management

Share data and video from a secure, cloud-based platform.
- Connect the entire team
- Cultivate consistency
- Connect your league

Hat-TriQ

by Stretch On Sense

The magic of Hat-TriQ

Hat-TriQ is a unique analysis platform assisting elite teams with detailed player analysis. How the players practice, how the players play the game, and the player's well-being are all gathered on one platform. Focused on empowering decision-makers with data, Hat-TriQ gives you detailed player and team analysis so your organization can make decisions based on true facts. As we combine all systems into one, you will reach your full potential for your entire organization. **With Hat-TriQ, there are no silos. IT's one platform.**

Hat-TriQ – The sports analytics' Spotify

Before Spotify, iTunes, and even Napster, we listened to music on a CD player.

If you wanted to listen to Bruce Springsteen's Born To Run, the song, not the whole album, you needed to buy that CD. If you wanted to listen to Glory Days from the album Born in the U.S.A., you needed to buy a second album from Bruce Springsteen. In that case, you now have 20 songs, but you only wanted to listen to 2 songs. You have too much music for your appetite and desire, which you also can compare to too much unfiltered data.

With Spotify, you have one platform with all artists, bands, and songs from the past, and what's fresh today and what will be cool and fresh tomorrow. You can choose one song from one album and quickly switch to a new artist and song. You can access all the music in the world and easily filter and listen to what you like.

With Hat-TriQ, you have the same option. We offer you all the systems on the market into one platform, and we can easily be the filter of your choice. We customized Hat-TriQ for you as Spotify creates your playlists.

HAT-TRIQ

STRETCH ON SENSE

Qlik

49ing

49ing is an ice hockey data and analytics company. The Data Cockpit, 49ing's all in one platform, powers hockey development at all levels, professional and amateur, from the national teams to youth programs. In Switzerland, 49ing is the official data provider to the National League and Swiss League. Beyond players and teams, officials, PSOs, marketing and TV partners get their specific video clips and data for their use cases. 49ing's computer vision game tracking from video easily scales data generation and analysis across all levels.

49ing works closely with leagues to improve their operations and IT-ecosystem. Examples include state of the art coaches' challenge and other enhanced video based operations providing real-time access to all camera signals from the venues to anywhere in the world in real-time.

Some questions to hockey analytics experts

Mikhail Amelin, Hudl

Who are you and what is your connection to hockey analytics?

My name is Mikhail Amelin. I have worked with Instat for 8 years and now work at Hudl. Having been involved in the development of Instat for Hockey from the very start I have seen the increase in data and analytics usage across the sport and worked with many clubs on how they can use video and data to enhance player, team and club performance.

How do you use hockey analytics in your job?

I work with hockey clubs directly. Helping them enhance their scouting, player development and team communication using Hudl Instat's database of video and data. My fundamental goal is to help clubs improve productivity through enhancing their pre and post match workflows, making it easier and quicker to access data that is tied to video and open up their ability to boost the success of their club.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

We teach people on a global scale why advanced data and video is a must have in their workflow. So regardless of where a club is on their analytics journey, we are able to help guide their use of data and how this impacts the way their club operates. Ultimately this is done through quick and easy access to data which is simple to understand and use to communicate areas of opportunity.

What hockey question would you like hockey analytics to answer next?

The ability to measure complete player impact in the game would be a huge step. xG, CORSI and other metrics fail to answer this question for clubs.

Where is the hockey analytics field going? What do you envision for the next 10 years?

The need and desire to use data in hockey is there, inevitably there will continue to be more data at our disposal, allowing clubs to be more sophisticated with their analysis and continue to build on their team, player and club performance. However, I see the introduction of physical data changing the game. This will allow clubs to look at player development from a different angle, considering the player's physical condition and their impact on the game.

Miika Arponen, Ässät Pori

Who are you and what is your connection to hockey analytics?

Miika Arponen, I work as an analyst for Ässät Pori in the Finnish Liiga.

How do you use hockey analytics in your job?

They are 95% of my work, I provide and analyze numbers for the GM and the coaching staff.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

I usually produce just written/spoken reports and some tables and graphs, but have also created an own dashboard tool which the GM and the coaches have access to.

What hockey question would you like hockey analytics to answer next?

I think the Holy Grail of hockey analytics is still how to measure defense.

Which hockey analytics method/notion is the most important/influential in your job?

Tracking data which we have available for us in Finland.

Where is the hockey analytics field going? What do you envision for the next 10 years?

I think the tracking data will make things a lot different in coming years when it gets more widely adopted and the data itself would be more accessible for public.

What was your main take-away from LINHAC 2024?

Not to disrespect any of the speakers or panelists, but the biggest take-away by far from the conference is always the networking and mingling with other people from the industry.

Michael Elmer, KINEXON

Who are you and what is your connection to hockey analytics?

Michael Elmer, Sports Scientist (M.A.) and former IceHockey Goalie-, S&C-, Video- and Assistant-Coach. Currently Sports Science Consulting at KINEXON and responsible for all EU Indoor-Clients (Male). Due to my coaching-background, I have an strong connection to hockey analytics

How do you use hockey analytics in your job?

Analytics in Hockey is my daily business because I support several professional hockey teams here in Europe, everything related to tracking-data for performance and tactics.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

I am in frequent discussions with all of our clients. Within regular meetings or visits on-site, I share the latest trends within hockey but also within other indoor sports that are using our technology from KINEXON, Besides this, I have a strong network within the field of sports analytics.

What hockey question would you like hockey analytics to answer next?

When will it be possible to combine performance data together with tactical data to gain very useful insights for coaches?

Which hockey analytics method/notion is the most important/influential in your job?

Analytics of player speed zones pending on the location on the ice, the period and the game situation.

Where is the hockey analytics field going? What do you envision for the next 10 years?

Puh, this is an tough question ... I hope that hockey in general will “open up” for more advanced analytics - independent from the current results during a season - and further more important, these results have to find their way into the daily practice with the athletes.

What was your main take-away from LINHAC 2024?

All participants showed an great spirit to develop the hockey analytics further. LINHAC 2024 was also well organized.

Leo Girod and Thorsten Apel, SportContract

Who are you and what is your connection to hockey analytics?

We are SportContract, an all-in-one platform that provides video and analytics for ice hockey. Our platform helps analysts, general managers, and coaches make smarter decisions by combining game statistics, player performance data, and video analysis all in one place. With SportContract, users can easily access the information they need to understand the game better, develop their players, and build successful strategies. We offer everything needed to analyze hockey and gain a competitive edge.

How do you use hockey analytics in your job?

At SportContract, we use hockey analytics to look at how players and teams perform. We analyze different statistics such as how effective a player is during games, the quality of their shots, and how well a team controls the puck. This information helps teams make better decisions about which players to sign, how to create effective lineups, and how to plan for games.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

We know that explaining data can be challenging, so we make sure to present our findings in a way that is easy to understand. We use clear reports and visuals to show the key points. We also have meetings and training sessions to help coaches and GMs learn how to use these insights.

What hockey question would you like hockey analytics to answer next?

One question we would like hockey analytics to explore is: ”How can we better measure the impact of defensive plays that don’t show up in traditional stats, like takeaways, intercepted passes, or blocked shots, but still make a big difference in the game?” Understanding these defensive actions could help us see more clearly how players contribute to their teams.

Where is the hockey analytics field going? What do you envision for the next 10 years?

Hockey analytics is growing fast, especially with new technologies like machine learning, artificial intelligence, and player tracking. In the next 10 years, we expect to see more detailed data analysis, using real-time tracking to better understand player movements and decisions. We think analytics will become an even bigger part of how teams operate, affecting everything from scouting and training to game strategies.

Is there something else related to hockey analytics that you would like the LINHAC audience to know?

We want the LINHAC audience to know that hockey analytics is not just about the numbers. While data and stats are very important, the real value comes from understanding these insights and applying them in the right context. As analytics continue to grow, it's important to have collaboration between data experts, coaches, and players to fully use these tools and push the sport forward.

What was your main takeaway from LINHAC 2024?

Our main takeaway from LINHAC 2024 was seeing how much hockey analytics improves with new technology. The presentations showed how player tracking data is being used to better understand player movements and strategies during games. It was exciting to see how these advancements can help teams make better decisions and improve their performance on the ice.

Andreas Hänni, 49ing

Who are you and what is your connection to hockey analytics?

Andreas Hänni, former professional hockey player of 20 years and founder of 49ing. During my playing days, I was always very interested in analyzing and learning the game. After my career, I started 49ing with my brother Michael on our mission to grow the game harnessing data and technology. We do have a focus on the professional sport but want to really provide value to all possible use cases across all levels. Anything related to enhance the user experience around ice hockey.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

Translate Data into natural language and underly with video image is by far the most efficient in my experience.

What hockey question would you like hockey analytics to answer next?

A significant question I would like hockey analytics to address is how to more accurately quantify the impact of defensive actions that do not directly result in turnovers or blocks, such as positioning and pressure that lead to low-quality shots.

Which hockey analytics method/notion is the most important/influential in your job?

It is important to listen to data and not go find what confirms your opinions. Therefore, it is not up to the human, but up to the evidence to lead the way in the research to optimize performance. The question what contributes how much to winning is leading the way, methods and concepts follow. I think AI will bring dramatic improvements in exactly this: data based answers on open questions.

Where is the hockey analytics field going? What do you envision for the next 10 years?

The hockey analytics field is moving towards more granular and real-time data analysis. Over the next decade, I envision a shift towards AI as a key component as anywhere else in life.

Albin N Maelum, Stretch On Sense

Who are you and what is your connection to hockey analytics?

Stretch On Sense is a SportsTech company specializing in sports analytics. We've developed a cloud-based platform named Hat-TriQ, built on Qlik technology, which provides comprehensive team and player analysis by integrating various market systems into one unified platform.

How do you use hockey analytics in your job?

Hockey analytics is at the core of everything we do.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

We work closely with our clients, partners, and industry experts. This collaboration helps us stay ahead of the curve, ensuring that our solutions are constantly evolving and adaptable to the fast-paced demands of the industry.

What hockey question would you like hockey analytics to answer next?

We're particularly interested in AI and real-time analysis, specifically how advanced player tracking data can be leveraged to optimize team strategies and individual player performance during games in real-time.

Which hockey analytics method/notion is the most important/influential in your job?

Our work revolves around in-depth team and player analysis, spanning the entire organization, including scouting. We specialize in breaking down the numbers to provide elite teams with actionable insights through detailed hockey analytics.

Where is the hockey analytics field going? What do you envision for the next 10 years?

Hockey analytics is still in its early stages, but the field is poised for significant growth and innovation. Over the next decade, we expect an increase in the amount and quality of data available for analysis, leading to more precise and detailed measurements. The systems we use will become even more advanced, ushering in a new era of analytics that will transform how the game is played and analyzed. We have an exciting future ahead in this field.

What was your main take-away from LINHAC 2024?

The main takeaway from LINHAC 2024 was the significant progress in real-time sports analytics and the growing use of AI in decision-making. The conference emphasized how these technologies are already being applied in games, transforming team strategies and performance. It also highlighted the importance of collaboration between tech providers, analysts, and teams to ensure practical and impactful innovations.

Jan Morkes, Bílí Tygři Liberec

Who are you and what is your connection to hockey analytics?

My position is data analyst/scout and together with my brother David, who specialises in analytics/skill development, we form one of the largest analytics departments in European hockey, working for Bílí Tygři Liberec in the Czech Extraliga:)

How do you use hockey analytics in your job?

I don't draw a big contrast between "analytics" and traditional hockey knowledge. In every part of our work, we're trying to combine the best possible information and come up with coherent answers. And because European organisations are less crowded than US ones, you tend to find yourself in the middle of discussions about everything. Scouting, game performance, internal player development, tactics, general trends in hockey. A complex knowledge of "analytics" underpins it all.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

The key is to do your homework and communicate the conclusions of your research to managers/coaches/players. Explain what it means in hockey terms and how they can use the information in a practical way. Use the data if you have to, but only to illustrate the point. Nobody "cares" about complex statistical analysis.

For us, it's not just about going to analytics conferences like LINHAC, it's also about visiting hockey coaches' conferences, listening to players and coaches, knowing how they think and what their problems are. My brother David's role is focused on understanding the deepest details of skill development so that he can discuss all the issues directly with the players.

What hockey question would you like hockey analytics to answer next?

More and better defensive metrics would be nice. We need better tools to properly measure and visualise the context of most of the puck events we've already known of. And defence is a big part of that.

Which hockey analytics method/notion is the most important/influential in your job?

Probably the basic scientific idea that you can measure a lot of things, but you have to conceptualise and operationalise them properly.

Where is the hockey analytics field going? What do you envision for the next 10 years?

I think we're moving from the era of limited data to the era of abundant data. There are a lot of companies offering interesting services, and sports organisations will need people who can navigate all these sources and help teams find the comparative advantage, not just blindly trust every green or red number on some random dashboard.

Inside the organisations, I hope the use of data analytics will become as casual and boring as the use of video. Everyone will do it and understand the benefits.

What was your main take-away from LINHAC 2024?

We could talk about new technologies, wishes about AI or other technological solutions, but the presentation that resonated most with me was by Erik Lignell. He focused on how to communicate our message and really help others grow. As analysts, we spend most of our time discussing how to do the right calculations, how to answer questions more accurately, but if the analyst is the only person in the organisation who's got that calculation right, it's pretty much useless. We should think about how we develop other people, not just our models.

Is there something else related to hockey analytics that you would like the LINHAC audience know?

We watch the games. All the time!:)

Josh Pohlkamp-Hartt, Boston Bruins

Who are you and what is your connection to hockey analytics?

My name is Josh Pohlkamp-Hartt. I am the Associate Director of Analytics for the Boston Bruins. My primary role is as the lead data scientist for the team.

How do you use hockey analytics in your job?

We use analytics to help the Bruins make well informed decisions. Analytics for us is data science to describe and predict elements of hockey and software engineering to organize our data, allow our staff to interact with it and develop tools to make our processes for decisions more efficient. We support a variety of areas within hockey operations from on-ice strategy to drafting.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

We try to use as many channels and styles of communication as possible, from passive approaches like having data available to discover on a website to more active methods like recommendations through written reports or verbally in meetings. It is important to us to be skeptical and inquisitive of all opinions and beliefs held by us and others within our organization, to help us better understand how our sport works and how best we can support this process. From this, we find ourselves asking a lot of questions, some leading and others open ended, to create healthy discussion and growth of familiarity with our analytical insights.

What hockey question would you like hockey analytics to answer next?

To me the biggest open area is in the gap between subjective evaluation and measurable actions. Specifically, there are unsolved gaps in defensive attribution of on-ice impact (are they a good defender?), hockey intelligence (do they have good hockey sense?) and compete/desire (are they a competitive player?). Scouts/coaches do not have unified definitions for these concepts and these concepts are not directly measurable actions. How can we create a clear understanding of these concepts that can be reliably evaluated with minimal bias?

Which hockey analytics method/notion is the most important/influential in your job?

There is no silver bullet in analytics. Each of our partners on our team have different needs and our methods/notions change. Fundamentally, we are always looking for our work to be actionable/useful for our staff and well understood/measured by us. For example, with our coaches we want metrics in our strategy reports that relate to parts of the game they can impact with the players and are repeatable/stable. Telling them something like "we were dumped in on too often last game" may not be a useful stat in the regular season because we only have partial control of the other teams' actions and we are not likely to face the same strategy from the next team. While novelty is fun, it will not help a team win.

Where is the hockey analytics field going? What do you envision for the next 10 years?

Looking at other sports and how hockey has developed to date, we are currently moving incrementally 2 steps forward, 1 step back as we innovate, learn and reset but the sport is still dominated by hockey-educated decision makers (former players, agents, coaches, etc.) which slows the feedback loop. This should change to another time of rapid growth once data-centric teams find success in top leagues or tournaments. Like all sports, teams are looking to copy the perceived competitive advantages of other organizations and the successes of analytics centric teams will lead to our next explosion of growth. That should happen in the next 10 years, leading to team executives increasing the resources and influence of analytics around the NHL which will trickle down to all other leagues. I also think that we will see some change in data access with more high quality data being available to the public which should create more understanding and interest from hockey fans.

What was your main take-away from LINHAC 2024?

I was left with the impression that there are lots of motivated individuals working in the space but not enough institutional support. The growth from year to year of hockey analytics is incremental but the intelligence and effort of the community is almost limitless. We are not fostering the passion of our community members enough and allowing them to reach their full potential. That is how our sport grows exponentially. I think of support as providing data access, opportunities to learn (internships, mentorships) and financial support for research.

Is there something else related to hockey analytics that you would like the LINHAC audience know?

Keep up the good work! If you are ever running low on ideas, look to our mature sports analytics neighbors like soccer and basketball. One area I would love to see more work in for hockey is Causal Inference. We do not do enough to understand the mechanisms of events and their outcomes. Player evaluation will improve greatly once we can start to do that!

David Radke, Chicago Blackhawks

Who are you and what is your connection to hockey analytics?

David Radke, Senior Research Scientist for the Chicago Blackhawks

How do you use hockey analytics in your job?

My job revolves around addressing open problems in hockey through developing novel methods at the intersection of artificial intelligence and hockey analytics. I use offline data collected in hockey games (both event and tracking) to build models with a variety of targeted use cases for stakeholders across hockey operations.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

The findings from my work are communicated through self-serve platform and written media that are accessible across the hockey operations department. Furthermore, I make myself available for follow up questions or analyses of the results which helps support future iterations of any project.

What hockey question would you like hockey analytics to answer next?

I'd like hockey analytics to take a more holistic view of how to build teams. It isn't a question that will likely ever be answered fully, but more of a shift in how we think about how our models value players and groups of players.

Which hockey analytics method/notion is the most important/influential in your job?

The most important/influential methods in my job come from the fields of reinforcement learning and multiagent systems. Most of my work deals with understanding how decisions are made, the value of those decisions, and how multiple players may perform together in specific contexts.

Where is the hockey analytics field going? What do you envision for the next 10 years?

My hope is that the field of hockey analytics is integral to team building in 10 years. This includes all NHL teams having analytics groups that are integrated across all areas of hockey operations to provide value. Furthermore, I hope that hockey analytics research groups are more commonplace across both academia and industry.

Robin Schürmann, d-fine

Who are you and what is your connection to hockey analytics?

I am member of d-fine's Sports Analytics team mainly working around hockey analytics. We help our clients, including sports clubs and federations, to effectively utilize their data for various purposes such as match analysis, scouting, and more.

How can hockey analytics models be operationalized?

To effectively apply insights from custom-build analytics models in practice, it is essential to prioritize user-friendliness for coaching staff and scouts. This can be achieved by making the models for match analysis or scouting easily accessible and seamlessly integrated into their workflow. One effective approach is to embed these models within intuitive and interactive dashboards, ensuring a low threshold for adoption and utilization. By streamlining the process and providing user-friendly tools, the practical application of hockey analytics can be enhanced.

How do you communicate hockey analytics findings to your clients?

When communicating hockey analytics findings to clients, it is important to make the information tangible. One effective approach is to use absolute numbers. Instead of stating that a shot has an expected goal value of 0.14, I illustrate it by saying it would take seven of these shots to score one goal.

What was your main take-away from LINHAC 2024?

My main take-away from LINHAC 2024 was the diverse range of hockey analytics topics covered, including fan engagement, cognitive studies, match analysis, and scouting analytics. I am particularly intrigued to see how hockey analytics will progress in these areas and the potential synergies that may emerge between them.

Fredrik Sjöö, Onsite sport

What is your connection to hockey analytics?

None, but interesting to find ways to bring interesting analytics to the fans, i.e. fan-engagement.

How do you use hockey analytics in your job?

Not, but personally interested and doing some analysis myself for my floorball club.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

I don't.

What hockey question would you like hockey analytics to answer next?

What is measurable is one thing, but how to measure team-building and social attributes that maybe not the best stats player, but a player that gives value to the team and make everyone better?

Which hockey analytics method/notion is the most important/influential in your job?

None.

Where is the hockey analytics field going? What do you envision for the next 10 years?

You gonna replay sequences in the game, in first person view within the game (3D rendered).

What was your main take-away from LINHAC 2024?

There are many analytics companies competing on leagues and clubs. There will be consolidations and as in most businesses 1-3 market leading ones left over the years.

Is there something else related to hockey analytics that you would like the LINHAC audience know?

We can educate and entertain the fans even more, exposing advanced analytics in smart ways to the audience with new technology in realtime. Making watching sports more immersive!

Simo Teperi, Rauman Lukko

Who are you and what is your connection to hockey analytics?

Simo Teperi, data analyst for Liiga team Rauman Lukko, I've worked in that role for 6 years.

How do you use hockey analytics in your job?

That's all I basically do. I gather the data from various sources and then use/model it in a lot of ways.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

I usually have a sort of table or chart which I explain how to read while presenting it to the coach or the GM.

What hockey question would you like hockey analytics to answer next?

Tactical stuff, like what forecheck would be best for this situation etc.

Where is the hockey analytics field going? What do you envision for the next 10 years?

I would probably look at where sports like soccer are now(or 5 years ago) and that would be a decent goal. More use of the tracking data and better quality.

Erik Wilderoth, Färjestads BK

Who are you and what is your connection to hockey analytics?

Erik Wilderoth, Assistant GM and Analyst at Färjestads BK.

How do you use hockey analytics in your job?

Daily. It is my job to do it.

How do you communicate hockey analytics findings to your customers/viewers/players/coaches?

By reports, emails, meetings and presentations. Depending on the receiver the method is changing of course. As it is fairly integrated in our organization and the organization is fairly mature in hockey analytics it's just information. A bit to the puzzle. Another layer to the onion.

What hockey question would you like hockey analytics to answer next?

More on the tactical part would be nice. Understand more off-puck impacts from players.

Which hockey analytics method/notion is the most important/influential in your job?

The old saying - "A scatter plot says more than a thousand words".

Where is the hockey analytics field going? What do you envision for the next 10 years?

For SHL I see it as the journey as many companies have gone with Data Scientists. First, it grows as a department but later on integrates with each department. So we will not have a "data department" but rather one analyst working with coaches, one with the performance team and one with the front office. Naturally integrated in the organization.

What was your main take-away from LINHAC 2024?

That we are growing as a field but maybe not in the speed I would have predicted three years ago. The organizations needs to mature before we, as a community, take further steps. One big step that needs to be taken there is to make data public. Without people knocking on the doors with competence to make a splash right away, clubs won't hire at the speed they should.

The research track at LINHAC shows that a lot of progress is made within the area and the boundaries are pushed forward but there may be a discrepancy between team's knowledge/maturity and the things we can do with data. Smart people need to get hired. And should be.

Student competition papers

Navigating the Rink: Analyzing Zone Entry Sequences and Expected Threat in Ice Hockey

Priyansh Gupta[1], Vignesh Aswathanarayana[1], and Ragini Gunda[1]

Linköping University liu.se

Abstract. This research paper analyses the dynamics of offensive zone entry in ice hockey, employing an adapted expected threat (xT) metric. Evaluating pass and carry sequences, investigating the influence of lane choices on offensive potential and defensive vulnerability. The findings offer insights for optimizing entry strategies and enhancing team performance.

Keywords: Zone entry · Markov's chain analysis · Expected threat(xT).

1 Introduction

Ice hockey, renowned for its intense pace and complex strategies, revolves around critical moments that finishes in scoring goals. The outcome of a match depends on many factors, ranging from skills of individual players to cohesive strategies deployed by the team. Among these factors, we look closely at offensive zone entries-the strategies that enable a team to advance the puck beyond the blue line and into the opposition's defensive zone. Despite the acknowledged significance of zone entries, our objective is to explore deep into this aspect of the game

Our research takes inspiration from the studies of Nick Czuzoj & Shulman [1], who have previously explored related themes, with a particular focus on the implications of offensive zone entries in relation to player handedness. While we draw upon their findings, our study seeks to identify patterns and tendencies that can inform more offensive strategies, thereby enhancing a team's ability to capitalize on scoring opportunities.

2 Background

Our analysis of offensive zone entry was conducted utilizing the event data sourced from one hundred and fifty-six games played during 2023 season in the Swedish Hockey League (SHL). To facilitate a detailed spatial analysis of gameplay, the hockey rink is segmented into lanes; effectively forming nine longitudinal lanes, as shown in Fig 1. The central lane is designated as "Middle", while the remaining eight lanes are symmetrically distributed as "Left", "Left Middle", "Right Middle", "Right"; with two lanes allocated to each category respectively. The division allows for detailed examination of player movements and puck trajectories across the rink.

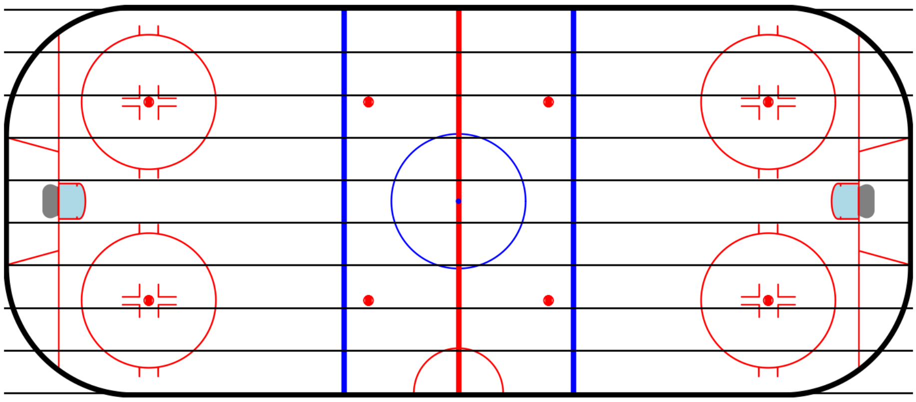


Fig. 1. Ice Hockey rink grid division

It is recognised that events such as the pass, carry, and dump-in facilitate players to enter the offensive zone. Our study focuses on passes and carries, which we categorize as controlled zone entries. This distinction aids in our understanding of how teams strategically move the puck over the blue line and into the opposition's territory, with an emphasis on the precision and intentionality behind these movements.

3 Algorithms

In this paper, we have two key methodologies to aid our analysis of zone entries, which are Markov's chain and entry sequence analysis.

3.1 Markov's Chain Analysis:

Markov chain analysis is a powerful probabilistic model, extensively used in sports analytics, adapted here from its original application in soccer metrics. This approach involves modelling the transitions between different events in the game, such as player positions and puck movements. By quantifying the probabilities of transitioning between different events, we can compute the xT for various scenarios on the ice. This allows us to assess the likelihood of scoring opportunities leading to specific offensive plays.

Expected Threat: The concept of xT, originating from Sarah Rudd's work in soccer[4], was further developed by Karun Singh [5], and adapted in context of ice hockey by Tim Keller[2]. Expected threat in ice hockey combines offensive threat (xT_{Off}) and defensive threat (xT_{Def}). This involves quantifying risk for each event, including shot attempts and preceding movements. Offensive threat is computed using probability distributions from binned statistics based on spatial coordinates, while defensive threat is assessed by mirroring offensive threat values and negating them. The overall xT for each event is derived by summing offensive and defensive threats, which is demonstrated in Fig 2.

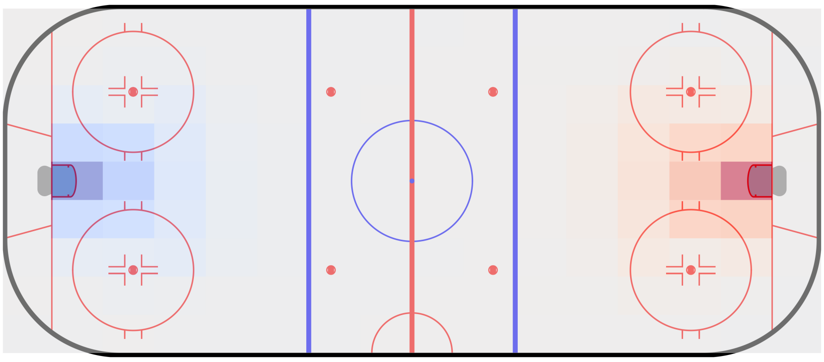


Fig. 2. Expected threat matrix on an Ice Hockey rink, blue colour represents the negative values (high defensive threat) while red colour represents the positive values (high offensive threat).

3.2 Entry Sequence Analysis:

In addition to Markov chain analysis, a custom method is employed to analyse entry sequences leading to controlled zone entries. This method tracks puck possession and events leading up to a successful entry. Key movements, puck movements, and event chains are identified to discover various entry strategies through the use of event tracking techniques.

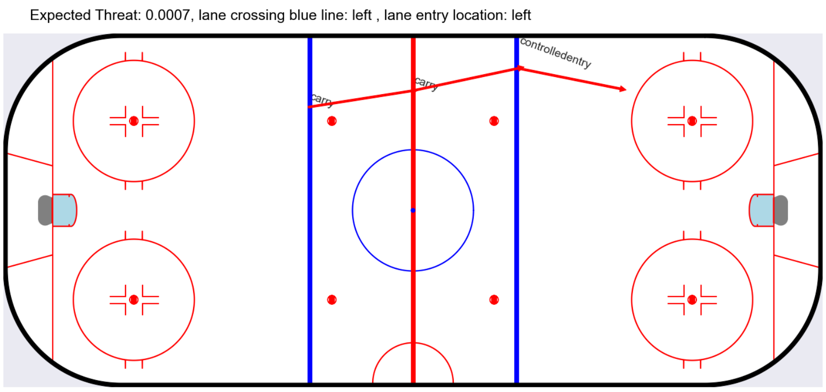


Fig. 3. An example sequence for the controlled entry into offensive zone

The methodology involves data preprocessing to isolate entry sequences. Through classification based on game events, entry attempts are categorized into passes, carries, and successful controlled entries. Spatial coordinates are used to analyse entry lanes and identify spatial distribution patterns. The impact of different strategies is assessed by associating the entry attempts with xT. One such example sequence is displayed in Fig 3.

4 Results

Entry sequence analysis is employed to determine the last few sequences before a successful zone entry. Typically, an offensive zone entry is finalized by the last action, but previous events such as passes and carries are also integral to the sequence leading up to a successful zone entry in the opposition defensive zone. Therefore, we have considered the last three events which leads up to a successful zone entry.

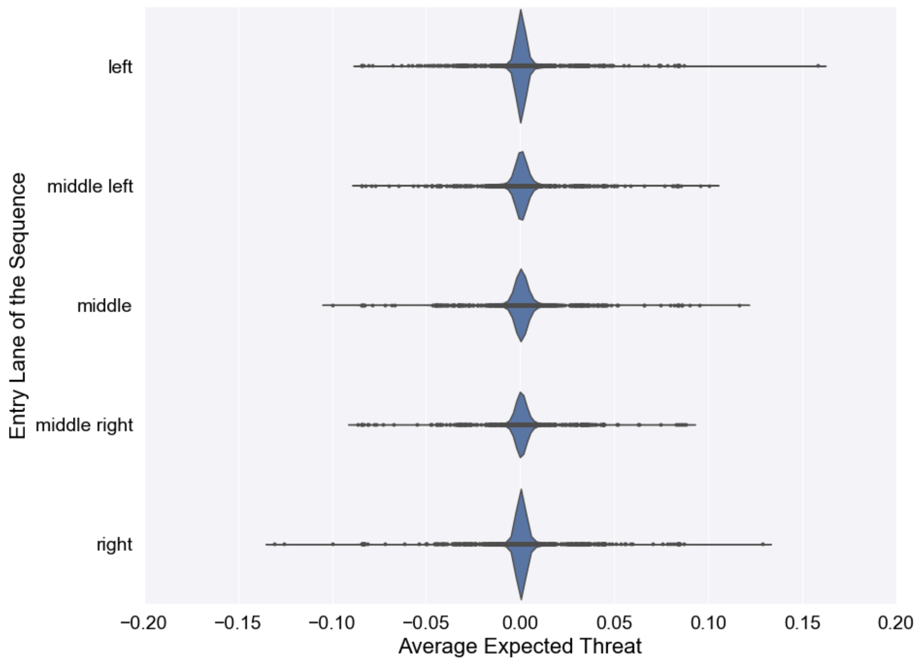


Fig. 4. Violin plot displaying the average expected threat for entry sequence group according to different lanes in the Ice hockey rink.

As the entry sequences are mixed, we didn't separate them based on different events like dump-in, pass, or carry. From Fig 4, we found that the left and right lanes exhibit high peak values. However, the middle lane also shows comparably high average xT within a sequence. Surprisingly, the right lane also exhibits the lowest value within a sequence, unlike the left lane. Overall, the plot provides evidence that the right and left lanes have the greatest number of sequences for offensive zone entry, thus indicating potential for the most extreme values.

Furthermore, we found an interesting trend when comparing the xT values for different lanes for two most occurring events: passes and carries while entering the offensive zone. Specifically, when entering the zone while passing the puck, the middle, middle left, and middle right lanes have the highest xT as represented in Table 1. Whereas, entering the zone while carrying the puck, the right, and left lanes have the highest xT. However, xT values in the case of carrying are negative, indicating that the defensive team poses a threat to the attacking team, thus increasing the chances of losing possession while carrying the puck in the offensive zone.

Table 1. Expected threat values for pass and carry along in accordance with the different lanes in descending order

Lanes	xT for pass
Middle	0.002503
Middle left	0.001492
Middle right	0.001092
Right	0.000332
Left	0.000048

Lanes	xT for carry
Right	-0.000411
Left	-0.000466
Middle left	-0.001649
Middle right	-0.002539
Middle	-0.009637

5 Summary

Our paper presents a thorough analysis of offensive zone entry in ice hockey using Markov chain and entry sequence analysis. Results reveal distinct characteristics of different entry lanes play a crucial role in shaping offensive strategies in the game. The left and right lanes are notable for their high offensive potential, offering wider angles that facilitate lateral movement and provide numerous passing options. The middle lane also emerges significant, providing a direct path to the net. Furthermore, considering xT values associated with different entry events, passing into the offensive zone through the middle left, middle, and middle right lanes consistently yield the highest xT values, indicating promising scoring opportunities. On the other hand, carrying the puck into the offensive zone through the left and right lanes risk losing possession. This suggests that precise puck movement in key areas of the offensive zone significantly increase the likelihood of creating scoring opportunities.

6 Future Works

Further investigation on sequence of events that leads to successful goal opportunities, risk of penalty can be done in future. We could gain a deeper understanding on key moments that leads to scoring opportunities with this transition. This would involve both offensive actions and defensive responses. The work of Anton Olivestam[3] on determining the sequence which lead to shots with high expected goal value, gives more insight about shots on goal.

Acknowledgments. We would like to thank David Sumpter and Aleksander Andrzejewski [6] for their work in the soccer for expected threat and possession chain to make it available publicly, which we modified in context of Ice Hockey. That part of code belongs to their copyrights, which is clearly marked in our published notebook as well. We do not take any credit for that part of code.

Link to GitHub Code: https://github.com/priyansh16/Linhac-2024

References

[1] Nick Czuzoj-Shulman. *SEAHAC 2022 — OZ Entry Lanes.* 2022. URL: https://www.flipsnack.com/75B8FBBBDC9/seahac-2022-oz-entry-lanes.html.
[2] Tim Keller. *Where did they get out? Evaluating zone exits using expected threat in hockey.* Linköping Hockey Analytics Conference. 2023.
[3] Anton Olivestam. et al. *On the Attack: Using Analytics to Unlock the Secrets of Successful Zone Entries in Hockey.* Linköping Hockey Analytics Conference. 2023.
[4] Sarah Rudd. *A Framework for Tactical Analysis and Individual Offensive Production Assessment in Soccer Using Markov Chains.* 2011. URL: https://docplayer.net/27070167-A-framework-for-tactical-analysis-and-individual-offensive-production-assessment-in-soccer-using-markov-chains.html.
[5] Karun Singh. *Introducing Expected Threat (XT).* 2018. URL: https://karun.in/blog/expected-threat.
[6] David Sumpter and Aleksander Andrzejewski. *Possession Chains — Soccermatics Documentation.* 2022. URL: https://soccermatics.readthedocs.io/en/latest/gallery/lesson4/plot_PossesionChain.html.

Decision Making in the Neutral Zone and its Impact on Possession Value

Ethan Parliament

Queens University, Kingston ON, Canada

1 Introduction

In the fast paced game of hockey, the neutral zone stands as a decisive area where the outcome of the game can hinge on split-second decisions. The swift and precise choices required in this zone can often determine the value of your possession, and ultimately, the game. To increase scoring opportunities, players must make the right choices when gaining possession in the neutral zone. This paper seeks to examine the impact that different actions taken in the neutral zone have on the value of a possession. By understanding the values added or removed by taking different actions, this study aims to equip players with the knowledge to make informed decisions when they gain possession.

2 Background Information

Creating strong possessions from the neutral zone is a crucial part of the game, as it often leads to extended possessions in the offensive zone, and more scoring opportunities. To assess the value of a possession, several metrics can be used. Cumulative expected goals (xG) serves as a strong indicator of possession strength. Possessions with multiple scoring chances and shots will have a higher cumulative xG than possessions with no shots. Another metric to value a possession is the duration of the possession. Longer possessions indicate control over the game, and can both create offensive chances and prevent the opposing team from scoring.

This paper aims to investigate the impact of the first decision made by the player with the puck when starting a possession in the neutral zone. The most common events were:

1. Passing the puck
2. Carrying the puck
3. Protecting the puck
4. Controlled Entry

A greater focus will be put on the differences between passing the puck, and carrying the puck, as those are by far the most common events.

3 Algorithms and Methods

3.1 The Dataset

The data provided for this competition by Sportlogiq included data from 156 games of the 2023-24 SHL season, with a total of 541,802 single events that were tracked. Each row represents a single event that occurred during the game, and includes data such as the team and player, event name, possession number, and the time and location of the event. The dataset also includes the xG for all shot attempts. These events were crucial for the continuation of this research. The coordinate system used in the dataset played a large role in this analysis. The circle at center ice is located at (0,0), with the defensive blueline located at x=-25 and the offensive blueline located at x=25. The boards along the neutral zone are located at y=42.5 and y=-42.5. This coordinate system will be used when showing plots of results.

3.2 Method

Before getting into the analysis, some pre-processing of the dataset had to be done to make the data easier to work with. The dataset was broken up into an array of individual possessions so that the value of each possession could be determined. Then, a filter was applied to only consider possessions that started in the neutral zone, which was determined using the location of the first recorded event.

Once the possessions were separated and filtered, a baseline of the metrics was created to compare to the different events. The average xG and average time of possessions was found for all event types for possessions starting in the neutral zone. Using this, it is possible to find the value gained or lost by making different decisions to start your possession.

To evaluate the cumulative xG of a possession, each single event in the possession was checked for its xG value, and summed up for the whole possession. This was then stored in a dictionary, with each possession starting event type (pass, carry…) as the keys, and an array containing the xG, and starting location of each possession as the values. This allowed comparison between different event types, as well as a comparison of the impact that location plays in these situations.

A similar process was followed to find the duration of each possession. The length of each possession was found by finding the difference in elapsed time from the start of the possession to the end of possession. Like with cumulative xG, this was stored in a dictionary along with the starting location of each possession, allowing for comparison of time of possession based on event type, and location on the ice.

4 Results

4.1 Comparing Baseline Metrics

Comparing the average xG and time of possession for different event types shows that the decision made by the player has a very clear impact on the value of the possession. Table 1 compares the impact that different event types have on the cumulative xG of a possession. It shows that by keeping the puck you can greatly increase the odds of scoring a goal. Passing the puck and protecting the puck will give a slight decrease in xG on average.

	All events	Pass	Carry	Puck Protect	Controlled Entry
Average xG	0.00558	0.00474	0.01395	0.00355	0.02038
% Change from all events	0	-15.05%	+150%	-36.38%	+265.23%

Table 1: Average xG by event type

Increasing time of possession shows a different strategy than trying to increase xG. Table 2 shows that by passing the puck you can increase the length of the possession. Carrying the puck or attempting an entry on the offensive zone will slightly decrease the length of the possession, and protecting the puck will greatly lower the length of the possession.

	All events	Pass	Carry	Puck Protect	Controlled Entry
Average Time of Possession (s)	5.0	6.45	4.57	3.57	4.20
Increase/ Decrease (s)	0	+1.45	-0.43	-1.43	-0.8

Table 2: Average time of possession by event type

The relationship between time of possession and xG can be seen in Figure 1. There is an inverse relationship between time of possession and xG. Possessions that are longer typically have lower xG. This is a tradeoff that teams are faced with, their strategy and game plan can dictate whether they want to control the game with time of possession, or go for higher xG, but shorter possessions.

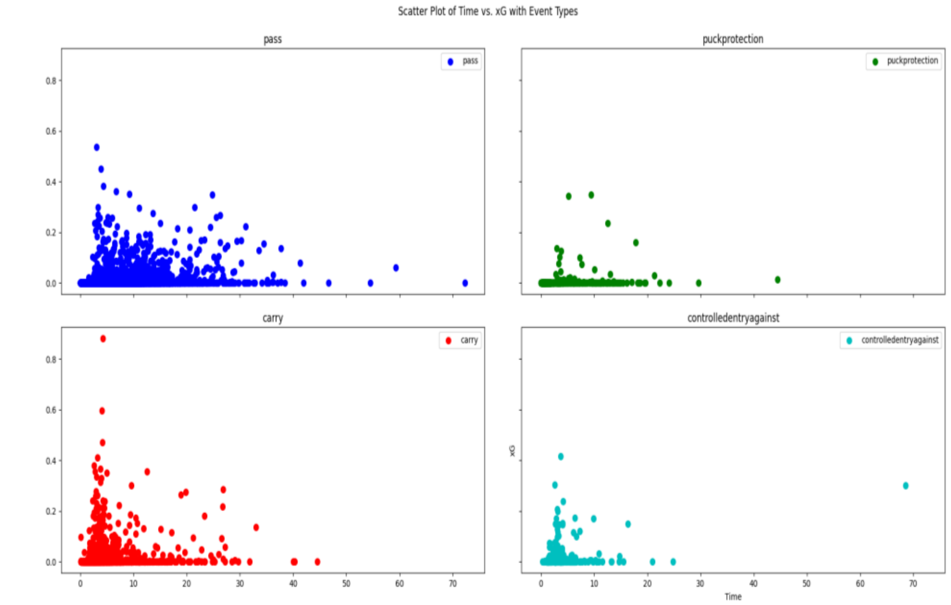


Figure 1: Comparison of time of possession vs xG for different event types

4.2 Location Based Metrics

When making these plays in a game, the player's location can play a role in deciding which read they should make. Figures 2,3,4,5 show average time of possession and average xg based on event type and location. The defensive blueline is at x=-25 and the offensive blueline is at x=25. Seen below in Figure 2 and Figure 3, the effect that location plays on time of possession is shown. In both cases, the starting location does not seem to play a significant role in changing the length of the possession.

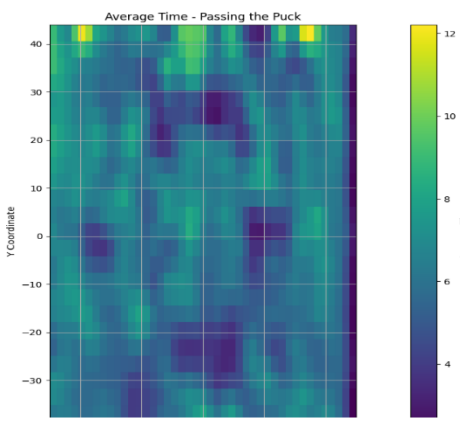


Figure 3: The average time of possession by location for passing the puck

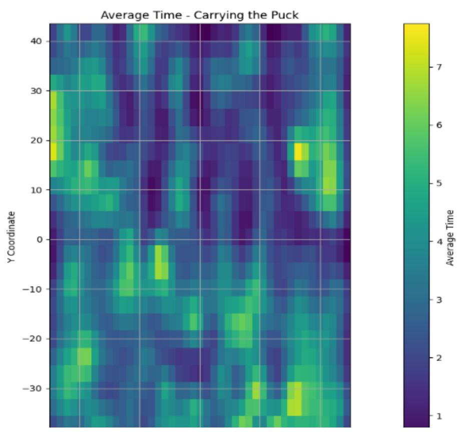


Figure 2: The average time of possession by location for carrying the puck

Unlike time of possession, the starting location of a possession does have an impact on the xG of a possession. Shown in Figure 4 and Figure 5, the effect of location does make a difference on whether the player should pass the puck or carry it. In the case of passing the puck, a greater advantage can be gained by passing the puck if the possession starts closer to the defensive blueline. The opposite can be said for carrying the puck, where it is more advantageous to carry the puck the closer you get to the opponents blueline. Knowing this, players can make better decisions by choosing to carry the puck if they are close to the opponents blueline, and to pass the puck if they are closer to the defensive blueline.

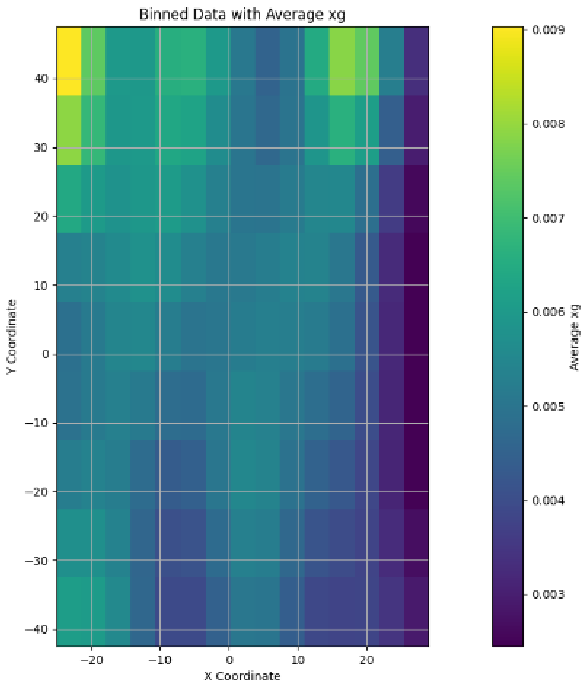


Figure 5: The effect of location on the average xG for passing the puck

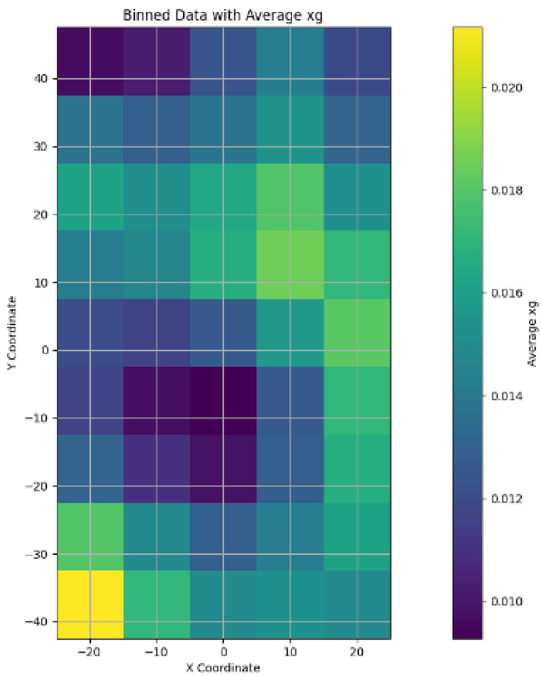


Figure 4: The effect of location on the average xG for carrying the puck

5 Summary of Results and Next Steps

The decision a player makes when starting a possession in the neutral zone significantly impacts the value of that possession. Given the fast-paced nature of hockey, players must quickly assess the situation based on various factors, such as the position of other players on the ice, and the game situation.

This analysis provides valuable insights to help players make informed decisions. For instance, when the puck is near the opponent's blue line, carrying it into the offensive zone is often the most advantageous option. If carrying isn't feasible, players should look to pass.

Moving forward, several avenues for further analysis present themselves. Exploring the direction of the next event, such as where to pass or skate, could enhance players' decision-making abilities, and guide them in the most effective direction. Additionally, investigating how a player's position or handedness influences decision-making could offer valuable insights. For example, forwards may excel at carrying the puck, while defensemen may be better suited to passing it.

By delving deeper into these areas, we can continue to refine our understanding of decision-making dynamics in the neutral zone and provide practical guidance for players and coaches alike.

6 Link to Code

https://github.com/eparly/Linhac2024

Beyond the Royal Road: A grid-based Approach to Identify Effective OZ Passes

Axel Rosendahl

Linköping University axero335@student.liu

Abstract. This study investigates the impact of crossing passes on scoring opportunities in hockey by expanding the traditional concept of the "royal road," an imaginary line that splits the rink into two halves. In this paper, the offensive zone is divided into the seven areas and analyzed sequences that contain passes that traverses these zones to understand. We tracked sequences and their outcomes across 156 games. The results highlight the effectiveness of certain zone transitions, particularly involving the slot, in increasing expected goals (xG) and contributing to team success. This paper offers insights into puck-zone transitions puck movement and its potential to improve scoring chances in hockey.

1 Introduction

Scoring goals in hockey can be done in multiple different ways. One particularly efficient method is to pass across an imaginary line, commonly referred to as the "royal road", which can be considered an imaginary line across the vertical centre of the rink. According to Valiquette, a pass crossing this line greatly increases the probability of scoring a goal, as the goaltender must adjust their position [1]. Instead of dividing the rink in two sides, this paper investigates if similar patterns can be found when the rink have been divided into seven different areas. Furthermore this paper also investigates how these passes relate to teams performance.

2 Background

Figure 1 shows how the offensive rink is divided. There are seven different zones, we are particularly interested in passes that traverses from zone to the other. The zones are referred to as (1) upper left, (2) upper middle, (3) upper right, (4) lower left, (5) slot, (6) lower right and (7) behind goal crease. This division is similar to Hellberg *et al,.* zone split of the rink in [2], but with the modification of accounting for which sides the passes or shots are coming from (left or right side).

3 Algorithm

The algorithm used in this paper can be described as a event sequence extraction algorithm. The algorithm is designed to identify two types of sequences

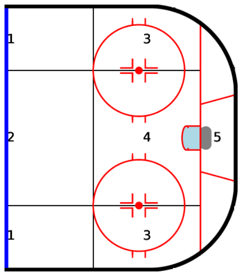


Fig. 1. The rink divided in predefined zones.

- **Sequence 1:**
  1. A *pass* from one of the predefined zones.
  2. A *reception* in a different predefined zone.
  3. A *shot*.
  4. The next event, checking if possession is maintained.
- **Sequence 2:**
  1. A *pass* from one of the predefined zones.
  2. An *assist* within a predefined zone.
  3. A *reception* in a different predefined zone.
  4. A *shot* on goal.
  5. Goal

These sequences are simple yet capable of capturing exactly what we seek to achieve; an analysis and quantification of what type of passes occur and which variants most successful. Furthermore this also used to identify correlation between the different passes and a teams success. The algorithm iterates through all rows in the datasets, and when a pass occurs in one of the zones of interest (referenced in Figure 1). If this pass is followed by an assist, a goal is guaranteed in this sequence since we require the next event should be a reception and a shot immediately afterwards. Alternatively if a the next event is a reception in another zone of interest it indicates a pass has changed zone, if the next event is a shot, we save the next event if the attacking team is still in possession. The result of this algorithm is two different type of sequences that capture the passes and puck movement we seek to analyze. This algorithm was chosen for its simplicity yet effective capability of capturing the interesting sequences.

4 Results

This section contains statistics about the different type of passes, xG values based on pass what zone the pass originated from and patterns that can be discovered from the sequences. Finally an analysis is done to investigate if any type of zone transition is particularly correlated to a teams success in terms of winning games.

4.1 Basic statistics the different type of passes and sequences

The data features 156 different games and the number of goals is 821. Out of the total 821 goals, 446 of these were found in sequence 2 sequences, in total around 9000 of sequences were identified meaning that around 5% of the sequences lead to a goal. Considering the the total size of the data set (156 games, and well over 500 000 rows), we manage to capture a significant amount of goals in relatively few events.

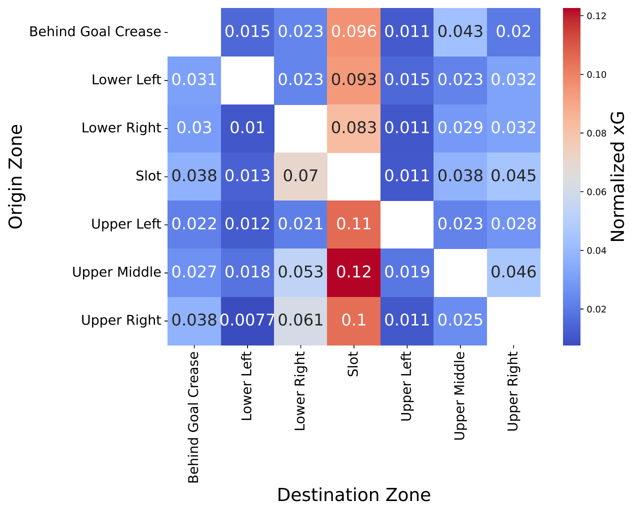


Fig. 2. Heat map off origin zone and destination zone.

Figure 2 shows the distribution and frequency of passes between the seven. Each cell in the heat map corresponds to the frequency of passes from an origin zone (rows) to a destination zone (columns). The color intensity in each cell reflects the normalized xG value where more red shades indicate higher normalized values and bluer shades indicate a lower normalized xG value. Recall that the normalized xG value is derived from dividing the accumulated xG-value for each pass-zone combination by the frequency of said pass-zone combination Not surprisingly, the slot is an important zone, receiving passes from all zones, leading to the highest xG-values, particularly the Upper Middle (0.122), Upper Right (0.105), and Upper Left (0.106). Significant activity has been observed behind the goal crease, highlighting its importance in starting plays. Notably, there is a strong link to the Lower Right area, with a xG value of 0.096. This suggests that teams might be using this area strategically to create scoring opportunities. Moreover, the analysis reveals a pattern of frequent lateral and diagonal passes, such as from the Lower Left to the Slot, which has a xG value of 0.093, and from the Upper Middle to the Upper Right, with a strength of 0.046. These insights not could be used to discover under-used combinations of the sequences.

Figure 3 displays a scatter plot where each dot represents an individual passing sequence within the dataset. The x-axis measures the total number of passes within each sequence, and the y-axis shows the corresponding total expected goals (xG) value. The plot includes a linear regression line, demonstrating a positive correlation: as the number of passes in a sequence increases, so does the xG value. This suggests that sequences involving more passes are likely to increase the likelihood of scoring, reflecting effective offensive play.

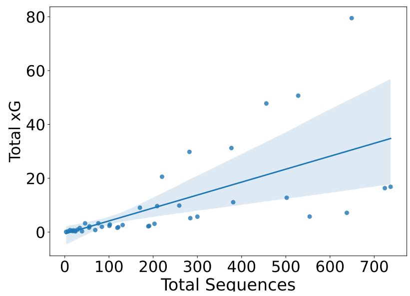


Fig. 3. Normalized xG vs Sequences.

4.2 Passing Zones correlation with team performance

In the earlier section we concluded that these types of passes are an effective way of scoring goals, the sequences are relatively infrequent but carry a majority of the goals in the total datasets. This is correlated with a teams success, from the dataset we identified team performance simply by looking at what team had won the most amount of games.

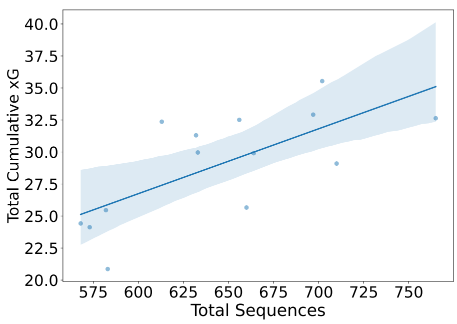


Fig. 4. Normalized xG vs Sequences per team.

Figure 4 shows a scatter plot where each dot is a team in the data set, and y-axis shows the normalized xG-value and y-axis shows the total amount of sequences. Furthermore we also have a linear regression line that suggests as the number of sequences that are transitioning from one zone to another, so does increases so the xG, which is a good measurement of overall team performance. Most of the teams follow the trend, however some points are outside of the confidence interval, teams that are out of the confidence interval suggests that it exists over-and-under performers in this aspect. This Figure provides visually that there is a relationship between having many sequences where a shot come from another zone than where it originated from is correlated to higher xG-values. This Figure share multiple characteristics with the similar plot based on total sequences, as seen in Figure 4.1.

Table 1. Performance Analysis in High xG Zones

Zone	Best Team ID	Best Team xG	Worst Team ID	Worst Team xG
Upper (M) to Slot	885	0.183	869	0.075
Behind to Slot	825	0.135	877	0.080
Upper (R) to Slot	885	0.141	792	0.060
Lower (R) to Slot	524	0.113	825	0.049
Upper (L) to Slot	634	0.166	686	0.056
Lower (L) to Slot	628	0.216	792	0.052
Lower (L) to Upper (M)	792	0.033	726	0.016
Upper (L) to Upper (M)	877	0.032	628	0.014

In Table 1 team 885 shows strong performance in generating higher xG from different field locations, conversely, Team 792 frequently appears as having the lowest xG, suggesting areas for improvement in their attacking plays. This table only shows a small subset of the available data. The data chosen was the highest xG transitions along with the top and bottom performers.

5 Overview and Discussion

The key findings in this paper is that it is established that are certain zones in the OZ that are more effective than the others, most of the findings are fairly trivial such as slot passes are effective in terms of xG, however when comparing the sequences in terms of performance wise in team we can get more novel and usable information.

6 Future work

This paper studied the effectiveness of transitioning from one zone to another in the OZ. A robust correlation was observed between these sequences and high team performance. I suggest future research to dig deeper into more team specific information such as presenting whether there are any significant deviations from the identified sequences with lots of action to and from the slot. This could lead to insights on what the teams should do and not do in the rink in terms of these zone transitioning passes. Lastly it would be interesting to validate the passing zones correlation with team performance with the results of the 156 games.

7 Code appendix

The can be found in this public repository

References

1. Valiquette, Steve `http://www.springer.com/lncs`, last accessed 2024/05/05
2. Olivestam, A., Rosendahl, A., Svens, H., Hellberg, L. On the Attack: Using Analytics to Unlock the Secrets of Successful Zone Entries in Hockey. In Proceedings of the Linköping Hockey Analytics Conference LINHAC 2023 Research Track." Linköping Hockey Analytics Conference LINHAC 2023. Linköping University Electronic Press, 2023

Pass or Shoot: The Great Dilemma in Ice Hockey 2v1 Situations*

Gunnar Samuelsson[1], Oscar Stolpe Östman[1], and Fabian Bergström[1]

Linköping University, liu.se

Abstract. This paper explores the factors and actions that lead to a high Expected Goals (XG) metric in two-versus-one scenarios in ice hockey, specifically examining the choice between passing and shooting. Through an analysis of average outcomes, we develop a preliminary model designed to inform players' decisions in these situations. While the model provides initial insights, its primary value lies in suggesting potential trends rather than prescribing definitive actions. Our findings offer a foundational understanding that may assist coaches and analysts in refining their strategies.

Keywords: 2on1 · 2v1 · xG.

1 Introduction

In this paper, we present an analysis of what actions that produce a high Expected Goals (xG) value in a two-versus-one situation in ice hockey. We specifically examine the impact of passing or shooting on the xG value in said situation. Two-versus-one situations in ice hockey are rare on a game basis, meaning that they happen only a couple of times per game. They are however high scoring opportunities, and scoring in such a situation could be the difference between winning and losing a game in the regular season, playoffs or finals. Determining the best plan of action for these situations is therefore crucial.

Previous work on the topic is limited and we cannot find any paper or report discussing this specific topic. The common belief is that in a two-versus-one situation, the defender should cover the pass, and let the goaltender take care of the awaiting shot from the puck carrier.

2 Background

A two-versus-one situation is defined as a situation where two attacking players are attacking against a single defender and the goaltender. In the provided dataset, this situation is registered by the defending team as an entry with event name "controlledentryagainst" and the type "2on1". In this situation, decisions has to be made quickly as the situation generally is over within five seconds.

* Supported by Linköping University

The player in control of the puck when the opportunity is registered has mainly three options. The player can carry the puck and shoot, try and dribble either the defender or goaltender or pass the puck to the teammate. Multiple passes between the attacking players are rare due to the short time window.

3 Algorithms

The work consists of a combination of scripts written in R. We will refer to the two primary scripts as "Script 1" and "Script 2".

3.1 Script 1

Script 1 sorts out all 2on1 events in the dataset and the 6 following events for each 2on1 sequence. These sequences are then divided into 2 categories. Sequences with a pass and sequences without a pass. These are refereed to as the "pass" category and the "shot" category. The sequences in the pass category are all the sequences which contain a pass in the 6 events following the registration of the 2on1 event and no shot prior to the pass. This means that if a pass was attempted, but has a failed outcome, meaning that the pass was not received by the teammate, it will be counted to the pass category. A situation like this will have 0 xG since no shot was taken. A situation with a successful pass and shot will be also be counted to the pass category with an xG determined by the shot. All shots have a listed xG in the dataset.

The pass category also has a subcategory which can be refereed to as the "completed pass" category. This category only consists of sequences which contain a completed pass with a following shot. The point of this category is to display the differences in xG for successful pass sequence compared to the average pass sequence.

Sequences counted to the shot category are sequences that contains a shot and no pass prior to the shot in the sequence. This sequence will also have an xG determined by the shot. If a pass was made after the shot, it will still be counted to the shot category. The important thing is that the shot was not made from a pass.

The choice to base the sequences on 6 events was made since the compiled game time for 6 events averages out to a time we determine appropriate for a 2on1 situation. This appropriate time window is about 5 seconds. We do not want to dismiss events within the 2on1 situation, but we also do not want to take into account events that happen after the 2on1 situation has passed.

The simple mathematical formula

$$\text{Average xG for sequence category} = \frac{\sum(\text{xG for all sequences})}{\text{Total number of sequences}}$$

is used to compare the xG for the different categories.

3.2 Script 2

The idea behind Script 2 was to calculate a weighted xG per pass and shot for each player in a 2v1 situation. This in order to determine which players that would benefit from either passing or shooting in a 2v1 situation. We only considered players that had appeared in a 2v1 situation in the dataset since they are of most relevance. For passing we first calculated each players passing accuracy, defined as all successful passes by the player divided by the players total number of attempted passes. Mathematically described as

$$\text{Passing accuracy} = \frac{\sum(\text{Successful passes})}{\text{Total number of passes}}.$$

This value was then used to define a new xG per pass for each player in 2v1 situations. The new xG was calculated by taking the total number of xG gained for all passes in a 2v1 situation in the entire dataset, divided by the total number of attempted passes in a 2v1 situation, weighted by a players passing accuracy. This weighted total number of attempted passes was calculated by taking the total number of successful passes in a 2v1 situation divided by a players passing success rate, therefore obtaining a weighted total number of attempted passes needed for each player to gain the total number of xG. The weighted attempted passes can be summarized by the formula

$$\text{Weighted attempted passes} = \frac{\sum(\text{Successful passes in 2v1})}{\text{Players passing success rate}}.$$

The new xG for a unique player can then be described by the formula

$$\text{Weighted xG per pass} = \frac{\sum(\text{xG for passing sequences})}{\text{Weighted attempted passes}}.$$

To calculate the weighted xG per shot in a 2v1 we compared each players actual goals versus the players expected goals gained on all shots taken. This generated a value to define a players shooting ability, given by the formula

$$\text{Player performance versus xG} = \frac{\text{Player goals}}{\sum(\text{xG for player shots})}.$$

This value was then multiplied by the average shooting xG given by Script 1 (Section 3.1) to calculate each players weighted shooting xG in a 2v1 situation. Worth noting is also the fact that passing sequences are twice as common as shooting sequences in the dataset.

4 Results

The results found by Script 1 can by summarized by Table 1. This data indicates that the highest average xG is achieved by shooting and not by attempting a pass. We can however also see that when passes are successful, a higher xG is achieved.

A sample of suggested unique player decisions based on the output from Script 2 can be seen in Table 2. The 3 players with the highest xG for shooting can be seen in Table 3. The 3 players with the highest xG for passing can be seen in Table 4. When counting the numbers of suggested shots and passes by Script 2, it can be concluded that 51.16% of players should shoot and 48.84% of players should pass.

Category/xG	**Pass**	**Completed pass**	**Shot**
Average xG	0.069	0.161	0.093

Table 1. xG for each sequence category as described in Section. 3.1

playerID	**xG shot**	**xG pass**	**Decision**
5038	0.155	0.092	shoot
23843	0.156	0.102	shoot
35236	0.056	0.091	pass
38465	0.119	0.102	shoot
79286	0.073	0.090	pass

Table 2. Sample of outputs from Script 2 showing suggested unique player decisions.

playerID	**xG shot**	**xG pass**	**Decision**
693747	0.226	0.097	shoot
580351	0.216	0.090	shoot
475158	0.206	0.089	shoot

Table 3. The top 3 players with the highest weighted shooting xG.

playerID	**xG shot**	**xG pass**	**Decision**
945327	0.047	0.107	pass
458311	0.045	0.102	pass
941262	0.073	0.102	pass

Table 4. The top 3 players with the highest weighted passing xG.

The average passing accuracy in general is 75% and in 2v1s 58%. If adjusting for this discrepancy to better represent the passing accuracy in 2v1s, the decision split becomes approximately 70% for shooting and 30% for passing.

4.1 Discussion

There are some things worth noting regarding how the calculations were made and the amount of data in the dataset that possibly could have effected the outcome of the results. The pass accuracy calculated to weight a players xG per pass in a 2v1 situation is slightly biased towards generating a higher xG per pass since passing accuracy in general does not represent a players ability to successfully pass in a 2v1. The weighted shooting xG calculation is however not biased since it is calculated compared to xG, which is based on situation. If the dataset was bigger the passing accuracy could be calculated in 2v1 situations leading to a more representative result.

4.2 Summary

The initial results from Script 1 suggest that going straight for the shot results in a higher xG on average in 2v1 situations. In situations where pass attempts were successful, an even higher xG was achieved. This tells us that passing in a 2v1 is high-risk-high-reward. Since passes in 2v1 situations occur twice as often as shots in 2v1s in the dataset, this tells us that players might have an intuition about this increase in xG for successful passes. The numbers does however suggest that they would be better of going for the shot on average.

The results from Script 2 indicate the same thing. This as the script suggest 51.16% of players should shoot and 48.84% of players should pass. This script also provides nuance to the problem as skilled passers possibly could leverage the increase in xG from successfully passing the puck in 2v1s. The results show that this seems to be the case for some players. Worth noting is also the fact that the most skilled shooters seem able to obtain a significantly higher xG for shooting than the most skilled passers can obtain for passing, again indicating that going straight for the shot might be an underrated tactic in 2v1 situations based on how often it is attempted in relation to going for the pass.

The points made together with the existing bias towards passing in our calculations mentioned in Section 4.1 leads us to conclude that shooting generally would be the preferred option for the majority of players.

5 Future ideas and improvements

A larger dataset with more 2v1 situations would strengthen the results as the passing and shooting accuracy then could be based on the situation on which it is being applied. This would make the underlying statistics more representative and the results more trustworthy. Future works could try to find passing patterns that possibly have lower risk while still leveraging the increase in xG that comes from a successful pass. Future works could also try to find patterns in events leading up to a 2v1. Taking into account who the receiving player is could also be an area of interest for future works.

6 Code appendix

The code used for this project can be found in the following GitHub Repository using the link: GitHub Repository

Clustering SHL teams using Offensive Zone Metrics*

Ali Raza Zaidi, Fauzan Lodhi , Xikai Wu, Ryan Cao, and David Awosoga

University of Waterloo, Waterloo ON N2L 3G1, CAN

Abstract. Advance scouting is essential to adequately prepare teams for upcoming contests. Traditional process involve spending copious amounts of time reviewing video and visually identifying key items. To expedite this process and more efficiently guide film study, this work analyzes offensive zone actions to identify and categorize patterns in strategy.

Keywords: hockey analytics · spatiotemporal data · sports statistics.

1 Introduction

An archetypal invasion-type game, ice hockey is built upon strategies to generate offensive opportunities and limit the quantity and quality of chances for the opposing team. Given the storied past of the sport, numerous approaches have been developed by teams to maximize their per-possession efficiency given their roster construction. In particular, insights into offensive zone strategies quantify how teams formulate an attack against their opponent, how these strategies vary within the league and what combinations of strategies are often used in conjunction with one another. The following analysis clusters teams by their offensive zone strategies within the Swedish Hockey League to analyze the different play styles within the league. Similar approaches have been utilized on a team level in other invasion sports such as soccer [1], but hockey-specific analyses either analyzed different facets of the game [2], the events the leading up to an offensive possession [3], or on individual player characteristics [4]. Therefore, utilizing a team-centered approach focusing on archetypes for offensive zone tendencies distinguishes this work from any predecessors and provides exciting grounds for analysis and exploration. The framework for this analysis identifies key events based on a combination of accumulated and derived metrics that reveal team tendencies on offence.

2 Background

This analysis is performed on event-level data from one hundred and fifty-six games from the Swedish Hockey League's (SHL) 2023-24 season. Each event tracked several features including the game ID, team ID, elapsed game time, location on the ice, team in possession and event type. Possessions in the offensive zone (o-zone) were tracked to determine the o-zone time, and events occurring

* Supported by the University of Waterloo Analytics Group for Games and Sports.

within this o-zone shift were measured to determine their frequency. Within the o-zone, the attacking team uses their possession to generate scoring opportunities. Certain offensive zone strategies emphasize high-risk plays that provide better-scoring chances but risk losing possession earlier. Other strategies prefer low-risk events where puck possession is paramount. Higher risk strategies will generally have a lower average offensive zone time, but will counteract that with higher quality and more dangerous shifts.

3 Methods and Algorithms

O-zone plays occur when a team makes a play within the opposing team's blue line until the end boards within the same half. Offensive zone shifts are defined as consecutive possessions occurring within the offensive zone. They start when an attacking team enters the offensive zone, or gains possession of the puck while already in the offensive zone, and finish when the defensive team gets possession of the puck, or when the puck leaves the offensive zone. The following events are marked as favourable outcomes for the offensive team and are tracked accordingly: **passes**, **shots**, **penalties drawn**, **goals scored**, **loose puck retrievals**, and **icings drawn**. We also derive secondary metrics based on the offensive zone time between the selected events. To justify the metrics chosen, passes and loose puck retrievals were tracked as they indicate control of the play, while shots (and goals) indicate the 'end goal'for the shift. Opposing team icings and penalties drawn indicate a successful o-zone strategy that forces the defensive team to commit a mistake. Cohesively, these paint a picture of a successful o-zone shift.

With this in mind, the idea is to use these metrics to cluster teams using unsupervised learning algorithms. Given the small number of teams (n=14) within our dataset, it was hypothesized that clustering algorithms like k-means may not be the best fit. Instead, the approach taken was that of `Gaussian Mixture Model`, which is formed from several Gaussian models describing the underlying domain [5]. The model generated was trained on the aggregate statistics, listed in Table 1. Additionally, the data was cleaned to remove non-regulation-time and non-even-strength events to ensure the data was unbiased, and it was then scaled.

Due to the small number of teams, we capped the number of clusters to a maximum of 6 to try and ensure that we wouldn't get many clusters of a single point. We also used Principal Component Analysis (PCA) to reduce the data's dimensionality, building visual intuition for how many clusters to expect [6]. Our PCA process was able to explain 96% of the variance with just 5 components. Then, based on manual inspection of the PCA graphs, present in Figure 1 below, it was determined that 5 clusters seemed to be the best case. We corroborated this using analysis of BIC scores [7], as in Figure 2 which when capped to 6 clusters, yielded its best result at $k = 5$ clusters, with the 5-cluster GMM almost exactly matching the PCA clusterings.

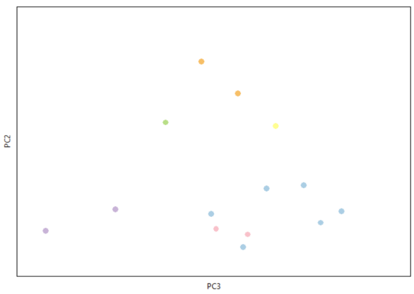


Fig. 1. PCA Clustering

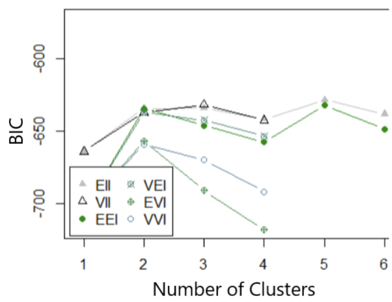


Fig. 2. BIC Clustering

4 Overview and Discussion of Findings

After training the `GMM` with 5 specified clusters, groupings emerged relating certain metrics to certain teams. We collected this data into a heatmap, labelled as Figure 3 below. Table 1 below highlights what each of the metrics highlighted in the analysis below are specifying. All derived metrics are reported in terms of their average values.

Table 1. Definitions of the Derived Metrics

Event	Description
time	time between shots
time_mod	time between shots*
pass	passes between shots
pass_mod	passes between shots*
ozone_time	length of the o-zone shift
max_xG	max xG generated within the o-zone shift
cmltv_xG	cumulative xG generated within the o-zone shift
goals	goals scored per the o-zone shift
passes	passes recorded within the o-zone shift
slotpasses	passes made to teammates in the 'slot'within the o-zone shift
fwdtime	length of time spent by forwards in the o-zone during the shift
dmentime	length of time spent by defencemen in the o-zone during the shift
penalties	penalties drawn within the o-zone shift
puckretrievals	successful puck retrievals within the o-zone shift
icings	opposing team icings committed within the o-zone shift

*shots without a pass in between are excluded

Based on the heatmap in Figure 3, we can see that cluster 1 generally has lower metrics than any of the other clusters, bar its goals and shots. Their low o-zone shift length and xG stats indicate that they struggle to maintain dangerous o-zone pressure. However, their fair number of shots indicates that they are taking the chances given. Thus this cluster has been labelled as an `Opportunistic Offence` given their propensity to capitalize on the few chances they get. Cluster 2, `Chip & Chase`, is more varied in terms of metrics, with below-average o-zone time, opposing team icings and defensemen o-zone time but above-average stats in xG, slot passes, and puck retrievals. The cluster struggles to keep the puck within the offensive zone, especially with the lack of a defence-

Mean Metric Values for Each Cluster

Metric (mean)	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
time	9.186	9.171	9.153	10.210	10.059
time_mod	11.091	11.115	11.462	12.297	11.822
pass	2.435	2.379	2.193	2.710	2.721
pass_mod	2.939	2.882	2.746	3.265	3.198
ozone_time	6.991	6.910	6.727	7.894	7.177
max_xG	0.045	0.051	0.051	0.051	0.044
cmltv_xG	0.047	0.053	0.052	0.053	0.045
goals	0.041	0.056	0.047	0.052	0.057

Mean Metric Values for Each Cluster

Metric (mean)	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
shots	1.123	1.119	1.086	1.124	1.095
passes	1.776	1.769	1.528	2.049	1.928
slotpasses	0.281	0.308	0.279	0.298	0.258
fwdtime	0.600	0.598	0.545	0.587	0.577
dmentime	0.325	0.336	0.365	0.346	0.366
penalties	0.008	0.009	0.014	0.008	0.005
puckretrievals	0.925	0.944	0.881	0.990	0.880
icings	0.010	0.006	0.016	0.014	0.006

Fig. 3. Identified Clusters

man anchoring the offence but also creates high-danger chances close to the net. Their puck retrieval stats also hint at the chip-and-chase nature of this cluster, where quality net-front opportunities are created through forechecking, at the cost of stability and longevity within the o-zone.

Cluster 3, `Off the Rush`, seems to have below-average metrics in most categories, except for xG, goals and penalties drawn. This would be indicative of a team relying on chances off the rush where opportunities would often be shorter in length, but with absurd xG, goal and penalties drawn metrics, given the possible odd-man rushes. Cluster 4 on the other hand, called `High-Pressure Cycling` has some of the highest metrics, particularly for o-zone time, slot passes, xG and puck retrievals, with their only below-average metric being penalties drawn along with a middling goals metric. This indicates a possession-heavy game, and their command of the middle of the ice gives them high-quality scoring opportunities using slot passes. They also seem to cycle the puck and force opposing teams to commit mistakes like icings. Their goals metric may indicate a lack of finishing ability. Finally, Cluster 5, `Point-driven Offence`, has above-average metrics in goals, defencemen time, and passes, but below-average metrics for xG, slot passes and shots. This indicates a perimeter-driven offence as they pass the puck often between the defencemen anchoring the offence but not often to the slot. Their point shots would be less frequent and would have a lower xG but the traffic created by the offence would be vital in scoring on more shots than expected.

Within a given cluster, certain attributes display a significant amount of variance, such as the ‘goals’ attribute. This highlights the clusters assembling based on the core playing style rather than team competency. Table 2 below lists the teams in each cluster. The anomalous nature of team 795 is highlighted as it is the only team playing its weak playstyle, but they make it work.

These results show that several different playstyles, along with their strengths and weaknesses, can be seen within the 2023-24 SHL season, though the degree of effectiveness of any given strategy will depend on the team employing them.

Table 2. Teams within each cluster

Cluster	1	2	3	4	5
Teams	795	855, 869, 814, 877, 792	524, 885	503, 686, 726	825, 634, 628

5 Conclusion and Future Steps

This analysis determined the 5 main types of team playstyles that exist within the 2023-24 SHL season, and where the teams fit into said clusters. These clusters can be used to analytically determine a team's preferred play style, and plan for it. With more games of data, we might be able to flesh out the model with more information, and with data from other leagues like the NHL which contain more teams, or even previous SHL seasons' worth of data, we would be able to cluster with more teams per label. Additionally, with player-tracking data, it may be possible to further define playing styles based on the whole team's positioning.

6 Code Access Links

The code used in this project can be accessed here: `https://github.com/AliRZ-02/linhac2024`

References

1. Coutinho, D., Gonçalves, B., Laakso, T., & Travassos, B.: Clustering ball possession duration according to players' role in football small-sided games. PloS one, 17(8), e0273460. (2022) `https://doi.org/10.1371/journal.pone.0273460`
2. Radke, D., Brecht, T., & Radke, D.: Identifying Completed Pass Types and Improving Passing Lane Models. Linköping Electronic Conference Proceedings, 71–86. (2022) `https://doi.org/10.3384/ecp191008`
3. Olivestam, A., Rosendahl, A., Hampus Svens, H., & Hellberg, L.: On the Attack: Using Analytics to Unlock the Secrets of Successful Zone Entries in Hockey. LINHAC 2023 (2023). `https://www.ida.liu.se/research/sportsanalytics/LINHAC/LINHAC23/papers/paper-students-LiU.pdf`
4. Stimson, R.: Identifying Playing Styles with Clustering. Hockey-Graphs (2017). `https://hockey-graphs.com/2017/04/04/identifying-player-types-with-clustering/`
5. Scikit-Learn: Gaussian Mixture Models `https://scikit-learn.org/stable/modules/mixture.html#gmm`
6. SciKit-Learn: Principal component analysis (PCA) `https://scikit-learn.org/stable/modules/decomposition.html#pca`
7. Scrucca, L., Fop, M., Murphy, T.B., Raftery, A.E.: mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models, 293 (2016) `https://journal.r-project.org/archive/2016/RJ-2016-021/RJ-2016-021.pdf`