Linköping Electronic Conference Proceedings Nr. 201

Edited by:
Tim Brecht
Niklas Carlsson
Mikael Vernblom
Patrick Lambrix

LINHAC

Proceedings of the
Linköping Hockey Analytics Conference
LINHAC 2023
Research Track

LINKÖPING UNIVERSITY

LINKÖPING

ISBN 978-91-8075-357-9 (Tryckt)

ISSN 1650-3686

eISSN 1650-3740

https://doi.org/10.3384/ecp201

Copyright: The Authors

Unless otherwise stated, is the content of this work licensed under a Creative Commons Attribution 4.0 International License.

Organizing Committee

Patrick Lambrix, Linköping University, Sweden (chair)
Niklas Carlsson, Linköping University, Sweden (co-chair)
Mikael Vernblom, Linköping Hockey Club, Sweden (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
Ulf Johansson, Jönköping University, Sweden
Patrick Lambrix, Linköping University, Sweden
Erik Lignell, Frölunda HC, Sweden
Joshua Pohlkamp-Hartt, Boston Bruins, USA
David Thomas Radke, University of Waterloo, Canada
Michael Schuckers, Statistical Sports Consulting and St. Lawrence University, USA
Oliver Schulte, Simon Fraser University, Canada
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

Silver industry collaborator

Hudl

Silver industry collaborator

Stretch On Sense AB

Silver industry collaborator

Sidelinesports

Bronze industry collaborator

49ing

Preface

LINHAC 2023 took place June 7-9, 2023, and was organized by Linköping University (Patrick Lambrix and Niklas Carlsson) and Linköping Hockey Club (Mikael Vernblom). 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.

The program included invited research talks by Frans Murto from Wisehockey on possession value models and Andrew C Thomas from SportsMEDIA Technology on puck and player tracking in near-real time. Further, four papers were selected for presentation in the research track of LINHAC. The program chairs for the paper selection committee were Tim Brecht and Niklas Carlsson.

In addition to the research track, Andreas Hofmann from Hudl and Thomas Krauskopf from Lausanne HC and the German national team talked about tailored and visualized feedback. Freddie Sjögren from Malmö Redhawks discussed perfomance analytics, while Johan Andersson talked about his experience with video analytics for the Swedish national team. Finally, Mike Kelly from NHL Network discussed the practical application of data in hockey and differences between regular season and playoff hockey in the NHL.

Further, there were six panel discussions moderated by Mike Helber and Niklas Carlsson. Two panels were made up of analysts, one with members from different SHL teams (Zack Ellentahl from Rögle BK, Erik Lignell from Frölunda Hockey Club, and Erik Wilderoth from Färjestad BK) and one with members from clubs of different leagues (Petter Carnbro from Leksands IF, Thomas Krauskopf from Lausanne HC and the German national team, Jan Morkes from Bílí Tygři Liberec and the Czech national team, Josh Pohlkamp-Hartt from the Boston Bruins, and David Radke from the Chicago Blackhawks). Two industry panels discussed the state of the art and future of hockey analytics where the second panel had an additional focus on data integration (Thorsten Apel from Sportcontract, Meghan Chayka from Stathletes, Miska Kuusisto from Wisehockey, and Morgan Zeba from Spiideo; Lance Du'Lac from Hudl, Andreas Hofmann from Hudl, Albin N Maelum from Stretch on Sense, Sean Tierney from Sportlogiq, and Freddie Sjögren from Malmö Redhawks). Goaltender analytics was discussed in a panel with Thomas Magnusson (Swedish Ice Hockey Federation), William Rahm (SCL Tigers), Maciej Szwoch (Färjestad BK), Sean Tierney (Sportlogiq), and Mikael Vernblom (Linköping Hockey Club). In the final panel, coaches and GMs discussed the use of analytics in their work (Adam Albelin from the Swedish Ice Hockey Federation, Patrik Hall from Växjö Lakers, Jeff Jakobs from Linköping Hockey Club, and Tomas Montén from HV71).

Our industry collaborators presented their products: Hudl, Stretch on Sense, Sideline Sports, and 49ing.

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. These research track proceedings include an invited paper and the papers from the research track. A companion book includes additionally contributions from industry, the student competition papers, as well as insights from contributors to LINHAC regarding their experience with hockey analytics and thoughts about its future.

We thank our moderator Mike Helber, our conference service TM Event, and the members of our local organization committee Mina Abd Nikooie Pour, Sijin Cheng, Huanyu Li, Ying Li, Gurjot Sing, Chunyan Wang, Jenny Rydén, 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 silver (Hudl, Stretch on Sense, Sideline Sports) and bronze (49ing) industry collaborators.

September 2023

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

Contents

Invited talk

Towards a real-time possession value framework in ice hockey

Frans Murto

Tampere University
frans.murto@tuni.fi

Abstract. Measuring the individual performance of players is an important task in sports analytics. Traditional statistics-based approaches for evaluating hockey players fail to account for context and long-term impact. Recent advances in data gathering have enabled valuing possessions and actions directly to address these issues. This talk describes the implementation of the first real-time possession value framework for ice hockey.

Keywords: ice hockey · possession value · tracking data

1 Introduction

Being able to objectively quantify performance in ice hockey has important implications for player evaluation and acquisition. Traditional goal- and shot-based statistics are problematic in this regard as they ignore the impact of preceding plays and do not differentiate between situational contexts. Advanced metrics derived from expected goals address the latter issue by incorporating information on how dangerous the shots a team takes or faces, but are unable to directly measure the value of non-shot actions and fail to consider situations where no shot is taken.

With the increasing availability of high-granularity event data streams and tracking data in recent years, various approaches in different team sports have been proposed for directly modeling the value of individual actions or possessions [1, 2, 4]. Related work in ice hockey has been scarce [5, 6], however, owing to a lack of the aforementioned data at least in the public domain. Our work intends to bridge this gap by proposing and implementing the first hockey-specific framework for deriving the total value of any given possession in real time, as well as valuing the risk and reward of individual actions separately.

2 Methodology

Estimating the value of possessions can be framed as a Markov decision process (MDP) where the possible actions a player in possession of the puck can take is represented by the discrete set A for all possible match states S. The actions we consider to be part of this action space are shots, passes, moving with the puck, dump-ins and dump-outs. Each action can be further separated depending on their outcome: whether a shot results in a goal, a pass reaches its intended target, a player maintains possession while moving with the puck, a dump-out is recovered by a teammate or a dump-out successfully exits the zone.

Players can be assumed to perform actions that intend to increase the probability of scoring for their team and decrease the probability of conceding a goal. Some actions that are valuable offensively, however, are inherently more risky despite their greater reward. To represent this trade-off we model the return of an action separately for both outcomes with success states yielding the probability of a team in possession scoring and failure states the probability of a team in possession conceding. We assume that only successful actions have positive returns and unsuccessful actions have negative returns, as even though some failed actions may lead to a positive outcome (e.g. a missed pass is received by another teammate in a relatively good position), from a modeling standpoint we want to consider only the intended target.

A common approach for representing the immediate and future impact of an action has been to use either a time window or a fixed number of future actions to assign positive labels for actions that end up affecting the score and negative labels for actions that do not [3, 4]. We experiment with different approaches and observe in our data that using a fixed window of eight seconds from the start of an action provides the best balance between short- and long-term return. When modeling reward actions we label actions that lead to the team in possession scoring a goal in the next eight seconds as positive, and when modeling risk we label actions that lead to conceding a goal in the next eight seconds as positive.

To formalize this we follow the definition of Fernández et al., where the value of a possession P_t is taken to be the total expectation of all actions in a given state [4]. The probability to take action a and its expected value are learned from $\mathbf{X}_t$, which is the feature vector representation of state s derived from a tracking data snapshot at time t.

$$\mathbb{E}[P_t] = \sum_{a \in A} \Big[\mathbb{E}[A = a|\mathbf{X}_t] \quad \mathbb{P}(A = a|\mathbf{X}_t) \Big] \tag{1}$$

As the outcome of successful and unsuccessful actions is modeled separately for all actions except shots (which we assign a fixed value of zero risk due to lack of a true failure condition), the expectation of an action can be generally decomposed as the difference between its expected reward and risk and how likely it is to succeed or fail. Because we assume a single end location for moves, the expected value of a move action follows this formulation exactly.

$$\begin{aligned} \mathbb{E}[A = \text{Move}|\mathbf{X}_t] = \mathbb{E}[A = \text{Move}_{Success}|\mathbf{X}_t] \quad \mathbb{P}(A = \text{Move}_{Success}|\mathbf{X}_t) \\ - \mathbb{E}[A = \text{Move}_{Failure}|\mathbf{X}_t] \quad \mathbb{P}(A = \text{Move}_{Failure}|\mathbf{X}_t) \end{aligned} \tag{2}$$

Because any teammate excluding goalies can be considered as the possible receiver of a pass, we take the expected value of a pass action to be the total expectation of all possible passes. We define an additional transition probability

$\mathbb{P}(R_t|\mathbf{X}_t)$ for all receivers R_t to represent how likely player r becomes the receiver of a pass.

$$\mathbb{E}[A = \text{Pass}|\mathbf{X}_t] = \sum_{r \in R} \Big[\mathbb{E}[A = \text{Pass}, R_t = r|\mathbf{X}_t] \quad \mathbb{P}(R_t = r|\mathbf{X}_t) \Big] \tag{3}$$

For dump-ins we consider all players except the goalie of the team in possession to be able to recover the puck. As dump-ins do not always have an intended receiver, but a general location instead, we model their success probability as $\mathbb{P}(A, R_t|\mathbf{X}_t)$ to represent how likely player r is to recover the dumped-in puck first. We take the expected value of a dump-in action to be the difference between the total expectation of all teammates T and all opponents O that can recover a dump-in.

$$\begin{aligned} \mathbb{E}[A = \text{DumpIn}|\mathbf{X}_t] = \\ \sum_{r \in T} \Big[\mathbb{E}[A = \text{DumpIn}, R_t = r|\mathbf{X}_t] \quad \mathbb{P}(A = \text{DumpIn}, R_t = r|\mathbf{X}_t) \Big] \\ - \sum_{r \in O} \Big[\mathbb{E}[A = \text{DumpIn}, R_t = r|\mathbf{X}_t] \quad \mathbb{P}(A = \text{DumpIn}, R_t = r|\mathbf{X}_t) \Big] \end{aligned} \tag{4}$$

For dump-outs we consider all opponents except the goalie to be able to intercept the puck in the offensive zone. We model $\mathbb{P}(A, I_t|\mathbf{X}_t)$ to represent how likely opponent i is to intercept the dump-out and use the total interception probability to determine the success probability. The expected value of a dump-out action is taken to be the difference between expected reward from successfully exiting the zone and the total expectation of risk for each possible interceptor.

$$\mathbb{P}(A = \text{DumpOut}_{Success}|\mathbf{X}_t) = 1 - \sum_{i \in I} \mathbb{P}(A = \text{DumpOut}, I_t = i|\mathbf{X}_t) \tag{5}$$

$$\begin{aligned} \mathbb{E}[A = \text{DumpOut}|\mathbf{X}_t] = \\ \mathbb{E}[A = \text{DumpOut}_{Success}|\mathbf{X}_t] \quad \mathbb{P}(A = \text{DumpOut}_{Success}|\mathbf{X}_t) \\ - \sum_{i \in I} \Big[\mathbb{E}[A = \text{DumpOut}, I_t = i|\mathbf{X}_t] \quad \mathbb{P}(A = \text{DumpOut}, I_t = i|\mathbf{X}_t) \Big] \end{aligned} \tag{6}$$

3 Data & Modeling

We use event and tracking data from the Liiga regular season matches 2020-21 and 2021-22 provided by Wisehockey. The player and puck tracking data is gathered using an indoor positioning system and sampled at a frequency of 20 Hz. The tracking snapshots are automatically synchronized by the system to align with the start timestamps of the events. We split 75% of the 872 matches in our data to train our models and use the remaining 25% as the test set. Events where the team whose scoring probability is being modeled faces an empty net are omitted due to the dynamics of scoring changing significantly in such situations. An overview of the events is presented in Table 1.

Table 1. Event data counts for the Liiga seasons 2020-21 and 2021-22.

Event type	Success	Total	Training	Test
Pass	79.87%	535,449	401,583	133,866
Move	95.35%	899,809	674,851	224,958
Shot	5.29%	78,141	58,608	19,533
Dump-in	39.67%	86,579	64,931	21,648
Dump-out	69.88%	48,836	36,630	12,206

Because the system provides situations where players move with the puck as continuous sequences, we split the puck controls into discrete one-second length actions. We assume that moves have one possible end location and set this as the player's position one second into the future based on their velocity at the time of the event. For unsuccessful passes we determine the intended receiver using a nearest-neighbor approach. We take the direction and velocity of the known trajectory of the pass and project it forward from the starting position of the pass at different time steps. The intended receiver is then chosen as the player who is closest to the projected end positions most frequently.

Each model uses positional features engineered from tracking data like location, velocity, direction, distance to goal, angle to goal, distance to puck and angle to puck for the player in possession of the puck and the target of the action (if applicable). To represent the local context of an action these features are calculated for the closest teammate and opponent of the player and the target. We also adapt the pressure model of Andrienko et al. to a hockey context and to consider the velocity of the players [7]. For the global context of an action we derive features using hierarchical clustering with two clusters on the players' locations to represent the tactical structure of both teams.

For the pass and shot expectation models we calculate features relating to the positioning of the goalie and how much of the net the goalie has to cover to make a save against the shooter or pass target. Pass expectation models also include information about the crowdedness and width of the pass lane between the passer and the target. The dump-in models have additional features based on the area where the puck is estimated to be played in for a given target. We use a simple convolutional neural network that takes raw tracking data snapshots as its input to estimate this end location. Finally, we include game and score state features in the action probability model to represent how teams adjust their playing style depending on the overall match situation.

We train our models using XGBoost[8] and optimize hyperparameters for tree depth and various regularization parameters using Bayesian optimization with cross-validation. To prevent overfitting into particular feature combinations we use moderate to high regularization parameter bounds for each model. As many of the learning tasks for our models consist of imbalanced data, we use a low max delta step parameter to ensure well-calibrated posterior probabilities. Based on domain knowledge we enforce monotonicity constraints on some features like distance to goal as this improves predictive performance and helps in dealing with outliers like goals scored from the neutral zone. Model metrics on the test set are presented in Table 2. We summarize the calibration of the models using expected calibration error (ECE), which takes the weighted average of the difference between binned output probabilities and data points.

Table 2. Model performance on the test set.

Model	Log loss	AUC	ECE
Action probability	0.186		
Pass success	0.423	0.803	0.037
Pass receiver	0.422	0.892	0.007
Pass scoring for	0.071	0.823	0.018
Pass scoring against	0.045	0.757	0.008
Move success	0.155	0.856	0.008
Move scoring for	0.053	0.791	0.012
Move scoring against	0.026	0.764	0.005
Dump-in success	0.165	0.873	0.006
Dump-in scoring for	0.029	0.743	0.002
Dump-in scoring against	0.015	0.750	0.003
Dump-out success	0.207	0.858	0.008
Dump-out scoring for	0.019	0.735	0.002
Dump-out scoring against	0.032	0.755	0.005
Shot scoring for	0.174	0.827	0.015

4 Results

As we can determine the total value of any given possession P_t, we can use this to evaluate the impact of each action. The value of an action is then defined by taking the difference between total possession value at the start of the action P_{Start} and the total possession value at end of the action P_{End} for all actions except shots, for which the total value is defined as-is. We define actions that have a positive impact as progressive and divide actions that have a negative impact into three categories: regressive, lost and conceded. Regressive refers to how much value is lost through successful actions that decrease the probability of the team scoring, lost refers to how much value is lost through unsuccessful actions, and conceded refers to the value gained by the opponent through unsuccessful actions. To account for any possible error in the detection of the moment of reception for passes, dump-ins and dump-outs, we take a one-second window after the end timestamp of the action to determine its end value.

The value of move actions is calculated continuously during matches, so we take the cumulative sum of a puck control sequence as the total value of a move action. As the value of move actions is the most sensitive to how other players are positioned, and we do not want to punish the puck carrier for factors outside their control, we limit the amount of regressive and lost value between two consecutive puck controls by the average value of a move action in our data (approximately one goal scored per 100 moves) and divide this limit by the sampling rate used to generate tracking data snapshots.

Because there are no existing ground-truth labels for evaluating the quality of a player's actions, we compare the total estimated possession value of common in-game situations and how often they have led to a goal being scored in the Liiga playoffs 2022-23. A comparison of the best forwards and defencemen that shows how their actions have added and lost value is also presented. We use playoff matches as our out-of-sample prediction because there is a recognized difference in play styles between the regular season and the playoffs, where players are allowed by the referees to play a more physical and disruptive game. This is illustrated by the fact that in our training data the average number of goals scored per match is 5.3, while in the 2022-23 playoffs it was 4.6 goals. We surmise that the features engineered from tracking data provide enough context to help overcome this domain shift.

Table 3 displays the danger level between controlled zone entry types at even-strength. We take the danger of a zone entry to be average of the total possession value in a one-second window following the moment that the offensive blueline has been crossed. We denote a zone entry to have resulted in a goal using the same eight-second window as with our models. Our framework correctly identifies that breakaways and odd-man rushes are likely to result in more dangerous situations than entries where the team in possession has an equal or lower number of players involved compared to the defending team. The estimated danger level generally correlates well with how often the different entry types result in goals, though some types are affected by the natural variance of goal-scoring in relatively small sample sizes. Based on observing individual entries, it is also fairly common for entry types to change a couple seconds after the entry is performed in the favor of the defending team as their forwards hustle to join the backcheck.

Table 3. Even-strength controlled entry danger by type in the Liiga 2022-23 playoffs.

Entry type	Count	Danger	Goal scored
1-on-0	18	12.7%	11.1%
2-on-1	24	7.4%	8.3%
1-on-1	43	5.4%	2.4%
3-on-2	70	4.6%	4.3%
2-on-2	146	4.3%	2.8%
1-on-2	90	3.6%	4.4%
3-on-3	473	3.3%	2.8%
2-on-3	310	3.1%	1.6%
1-on-3	101	2.6%	1.9%

Table 4 displays the danger level of successful even-strength passes by type. We take the danger level of a pass to be its progressive impact as defined above. Passes into the slot, which end in the area between the goal and the first hash marks of the faceoff circle, are intuitively the most dangerous with lateral (east-west) passes and behind-the-net passes inside the offensive zone following suit. We define entry passes as passes that facilitate a zone entry and outlet passes as passes that lead to a zone exit. Low-to-high and high-to-low passes are passes inside the offensive zone that do not cross the center line. It can be seen that the value added by these pass types closely follows the true scoring probabilities.

Table 4. Even-strength pass danger by type in the Liiga 2022-23 playoffs.

Pass type	Count	Danger	Goal scored
Slot	195	9.1%	9.2%
Lateral	994	3.6%	3.4%
Behind the net	226	3.3%	3.1%
Entry	611	2.3%	2.0%
Low-to-high	610	2.1%	1.8%
High-to-low	1142	1.4%	1.8%
Outlet	1631	1.1%	0.9%

Tables 5 and 6 represent the best-performing forwards and defencemen by total value added and lost at even-strength play. Passes, dump-ins and dump-outs have been consolidated together into a single value. The positive impact of an action indicates the progressive value added and the negative impact indicates the sum of regressive, lost and conceded value as defined before. Forwards generate more progressive value on average, but tend to lose more as well through attempting more ambitious and difficult actions. A greater proportion of the negative impact by defencemen is through conceded value, however, as their unsuccessful actions tend to occur closer to their defensive zone.

Table 5. Even-strength forward performance per 60 by possession value type in the Liiga 2022-23 playoffs.

Player	Team	Pass+	Pass−	Move+	Move−	Shot+	Total
Eemeli Suomi	Ilves	0.95	0.53	1.21	0.50	1.00	2.13
Anton Levtchi	Tappara	1.08	0.52	1.16	0.26	0.59	2.05
Joona Ikonen	Ilves	0.88	0.38	0.58	0.15	1.10	2.03
Waltteri Merelä	Tappara	0.48	0.39	0.93	0.34	1.20	1.88
Kristian Tanus	Tappara	0.91	0.49	0.91	0.24	0.73	1.82
Balázs Sebők	Ilves	1.04	0.81	1.31	0.23	0.50	1.81
Santeri Virtanen	Ilves	0.69	0.38	0.66	0.20	0.99	1.76
Matias Mäntykivi	Ilves	1.03	0.45	1.00	0.44	0.59	1.73
Niko Ojamäki	Tappara	0.75	0.35	0.44	0.15	0.97	1.66

Table 6. Even-strength defenceman performance per 60 by possession value type in the Liiga 2022-23 playoffs.

Player	Team	Pass+	Pass−	Move+	Move−	Shot+	Total
Les Lancaster	Ilves	1.06	0.43	0.33	0.10	0.82	1.68
Tarmo Reunanen	Lukko	0.76	0.28	0.68	0.20	0.58	1.54
Colby Sissons	KalPa	0.69	0.31	0.79	0.17	0.40	1.40
Valtteri Kemiläinen	Tappara	0.93	0.22	0.60	0.16	0.22	1.37
Maksim Matushkin	Tappara	0.69	0.20	0.70	0.21	0.33	1.31
Ben Thomas	Tappara	0.58	0.27	0.44	0.09	0.63	1.29
Leo Lööf	Ilves	0.82	0.32	0.62	0.21	0.15	1.06
Casimir Jürgens	Tappara	0.59	0.28	0.42	0.09	0.42	1.06
Thomas Grégoire	Lukko	0.71	0.28	0.43	0.12	0.22	0.96

5 Conclusions

In this talk we have shown that a real-time possession value framework can be implemented in an ice hockey context. The out-of-sample performance of the framework in typical in-game scenarios and events as well as in differentiating the value created by players in different positions matches both domain knowledge and the true underlying scoring probabilities. In the future it would be interesting to extend the framework to model defensive actions and the possibility of shots being indirect passes. Similarly, considering banked and rimmed passes separately would likely improve the performance of pass-related models [9]. Another promising avenue of research would be to use graph-convolutional neural networks with tracking data snapshots, which has been shown to improve model performance over tree-based models and remove the need for advanced feature engineering [10].

References

1. Sicilia, A., Pelechrinis, K., Goldsberry, K: DeepHoops: Evaluating micro-actions in basketball using deep feature representations of spatio-temporal data. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2096–2104. (2019)
2. Burke, B.: DeepQB: Deep learning with player tracking to quantify quarterback decision-making & performance. 2019 MIT Sloan Sports Analytics Conference (2019)
3. Decroos, T., Bransen, L., Van Haaren, J., Davis, J.: Actions speak louder than goals: Valuing player actions in soccer. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 1851–1861. (2019)
4. Fernández, J., Bornn, L., Cervone, D.: A framework for the fine-grained evaluation of the instantaneous expected value of soccer possessions. Machine Learning, **110**(6), 1389–1427 (2021)

5. 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)
6. Liu, G., Schulte, O., Poupart, P., Rudd, M., Javan, M.: Learning agent representations in ice hockey. In: Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2020 (2020)
7. Andrienko, G., Andrienko, N., Budziak, G: Visual analysis of pressure in football. Data Min Knowl Disc **31**, pp. 1793–1839 (2017)
8. Chen, T., Guestrin, C.: XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 785–794. (2016)
9. Radke, D., Brecht, T., Radke, D.: Identifying completed pass types and improving passing lane models. In: Proceedings of the Linköping Hockey Analytics Conference LINHAC 2022, pp. 71–86. (2022)
10. Stöckl, M., Seidl, T., Marley, D., Power, P.: Making offensive play predictable - using a graph convolutional neural network to understand defensive performance in soccer. 2021 MIT Sloan Sports Analytics Conference (2021)

Research papers

Professionalism & Leadership Development in Ice Hockey: Understanding Social Emotional Learning Experiences of Coaches in Atlantic Canada

Dr. Lynn LeVatte[1], Dr. Christina Phillips[2],
Dr. Kristin O'Rourke[1], and Shaun Ranni[3]

[1] Education Department, Cape Breton University, Nova Scotia, Canada (Lynn_levatte@cbu.ca, Kristin_orourke@cbu.ca)
[2] Department of Curriculum, Teaching & Learning University of Toronto, Ontario, Canada (c.phillips@mail.utoronto.ca)
[3] Department of Athletics, Cape Breton University, Nova Scotia, Canada (Shaun_ranni@cbu.ca)

Abstract. This qualitative research study investigated the Social Emotional Learning training experiences of ice hockey coaches in Atlantic Canada. Social Emotional Learning (SEL) and leadership in sport is an emergent field which has been gaining attention on a national level. The purpose of this study was to examine various aspects of SEL within coaching leadership training of Canadian ice hockey coaches in Atlantic Canada. Minor hockey coaches (n=8) were recruited to participate in semi-structured interviews. Five questions pertaining to hockey coaching background, leadership training, communication, and SEL training experiences were posed to participants. Interviews were offered both in-person or virtually as an option for convenience. Analysis of data suggested that clear expectations and effective communication with players and guardians were valuable aspects for relationship building. Limited professional development opportunities surrounding aspects of SEL were noted by participants, training provided was outdated in certain aspects, and current topics of inclusion, diversity, and culture. Future recommendations for continued study within the field of SEL within ice hockey are offered.

Keywords: Social Emotional Learning, ice hockey, coaching, professional development, leadership in sport, training

1 Introduction

Hockey in Canada has been conceptualized as a cultural truism and a way of life, with a connection so powerful and strong that it has united a vast nation from coast to coast [1]. The sport of hockey is embedded in Canada's national culture. Players, coaches, managers, and officials are key participants for the overall success and sustainability of this popular sport. To compare within a global perspective, there are approximately 1.8 million people who are registered hockey players, of which over one-third, or 631,295, live in Canada. There are 555,935 registered hockey players in the United States, 113,425 in the Czech Republic, 105,059 in Russia, 76,387 in Finland, and 63,901 in Sweden, rounding up the top six ranked countries [2]. The field of sport leadership has also emerged as a notable area of research and covers an increasingly diverse range of topics relevant to success in coaching youth athletes. The sport of ice hockey has grown to include many diverse populations and promotes diversity and inclusivity. While such perceptions of inclusivity have remained prevalent in many sports, recent critical events such as abuse scandals, racism, and bullying have negatively impacted the sport of ice hockey. The sport of ice hockey has witnessed growth within inclusivity and diversity aspects. In Canada, for example, the growth in female hockey has been substantial. Adams and Leavitt [3] reported that "the initiatives of women's sport leaders have led to greater recognition within the local, provincial and national governance structures and increased participation numbers". Exploring the 1980s, Canada saw an expansion of programs across the country to include opportunities for girls and women, and in 1982 a national championship was established [3]. As Hockey Canada [4] posited, within the 2009-2010 season, there were 85,624 girls and women registered as ice hockey participants, an exponential increase from the 8,146 participants reported two decades earlier in 1990.

Preparing to meet the coaching needs of diverse populations in the sport and essential training development through leadership is warranted. The International Ice Hockey Federation [2] explained that there are approximately 5,000 outdoor rinks and 3,300 indoor hockey arenas in Canada, and the only other country that has more than one thousand indoor arenas is the United States with 1,535. It can be determined that the sport of hockey has its greatest presence in Canada in terms of diversity growth, places, and facilities designated for ice hockey.

Urquhart et al. [5] explained that many definitions of coaching effectiveness mentioned winning in the professional context, while placing a greater emphasis on coaches developing athletes' confidence, competence, connection, and character. Lara-Bercial and Mallett [6] investigated characteristics of coaches and relationship to leadership. The findings of their study revealed coaches were characterized by 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 the Social Emotional Learning (SEL) training experiences of ice hockey coaches from the Atlantic Hockey Group (AHG). Within the past decade, the attention of SEL has broadened within society. As described by Liew and McTigue [7], educating the "whole child" became more prominent, thus enhancing teaching and coaching skills of professionals who work directly with youth. Within the expansive growth and popularity of diversity in Canadian hockey, leadership and SEL training to address this growth is imperative.

1.1 Impact of Social Emotional Learning (SEL)

Social Emotional Learning (SEL) can be explained as methods in which youth and children tend to learn, recognize, and manage emotions, develop positive relationships, behave ethically and responsibly care about others, make good decisions, and avoid negative behaviors [8]. It involves "teaching children to be self-aware, socially cognizant, able to make responsible decisions, and competent in self-management and relationship skills" [9]. In Figure 1, SEL general competencies are presented. These include self-awareness (e.g., identifying and recognizing emotions), social awareness (e.g., respect for others), responsible decision making (e.g., problem identification and situation analysis), self-management (e.g., self-motivation and discipline), and relationship management (e.g., communication, social engagement). These competencies parallel much of the sport-based research on life skills and psychosocial development of children and youth yet have largely remained isolated from the sport-based life skills development literature. SEL is critical for children and youth long-term success in and out of school [10]. Examining the intrapersonal characteristics for success, in-depth personal reflection, emotional intelligence, and a quest for continuous improvement have been instrumental within SEL research [11].

Furthermore, Elias [12] discussed SEL in sport and athletics, specifically investigating how SEL can be used to promote character development among athletes. 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 within itself to teach life lessons and prepare young people with values for lifelong learning [12].

SEL and behaviour in youth sport was recently investigated in academic literature. Research syntheses, systematic reviews, and meta-analyses support the development of SEL skills for promoting positive youth outcomes and reducing behavioral health challenges [14]. Youth who develop SEL skills can apply the "knowledge, skills, and attitudes necessary to understand and manage emotions, set and achieve positive goals, feel and show empathy for others, establish and maintain positive relationships, and make responsible decisions" [15]. CASEL cites many benefits of SEL skills, including improved attitudes, prosocial behavior, positive relationships, and academic performance [15].

This qualitative research study investigates the SEL training experiences of ice hockey coaches in Atlantic Canada. Professionalism and leadership in sport is an emergent field, which has been gaining attention on a National level [16]. Within the past decade, the attention and importance of SEL has broadened within society. As described by Liew et al. [17], educating the "whole child" became a more prominent approach within education pedagogy, and thus enhanced teaching and coaching skills of individuals who work directly in supporting youth in sport. Framed on the theory based upon Bronfenbrenner's ecological systems theory, which explains that environmental and contextual factors are necessary for understanding human development [18], this research project recognizes the potential and critical role that hockey coaches may play within the lives of youth.

Fig. 1: SEL framework by the Collaborative for Academic, Social, and Emotional Learning [13]

Education research aligns with this theory, as it points to the importance of community, teacher, parent, guardian recreation, schooling, and extra-curricular activities. These are the underpinnings for a holistic approach to education in which the interrelationships at all levels of the educational ecosystem prioritize human development alongside traditional education [19].

Recent research also suggests that ...the synergy between the practice of physical-sport activity together with physical and psychological health is a gradually growing interest area for education researchers [20]. Concepts that are characterized as constructs that are not identified with traditional indicators of cognitive capability or intellectual functioning and are often described under such terms as 21st-century skills, socio-emotional skills, character, or personality. Creating awareness of those skills can be beneficial for youth and coaches alike [20]. Participation in sports has also been related to a variety of social and emotional competencies and related skills that are correlated from extensive research and are essential to general success and well-being in school, work, and relationships [17]. Evidence suggests that youth with strong social and emotional 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 [21]. Previous research has determined that when enhancing youth social and emotional learning, one important factor is the ability and experiences of educators and coaches to engage and support learners [21]. Anderson-Butcher et al. [14] explained that behavioral and social skill development among youth is a growing concern. Training youth sport workers and community leaders within

SEL has the potential to positively engage youth [14]. Thus, this proposed research project will understand and determine existing levels and types of SEL training minor hockey coaches have received within their coaching training and explore potential need for future specific training.

2 Methodology

Data was collected for this study during the period of November 2022 to March 2023. Participants were recruited via an email sent from the project partner The Atlantic Hockey Group (AHG). The AHG was founded in 1989 by former NHL hockey player Charlie Bourgeois. It has become one of the most successful hockey training programs in Atlantic Canada. Through its many different hockey schools and leagues, the AHG works with over 10,000 hockey players each year. The team of experienced coaches come from the ranks of university, junior, high school, and minor hockey teams. The AHG is based in Moncton, New Brunswick, Canada, with programming offered in three Atlantic Canada provinces (Nova Scotia, Prince Edward Island, and New Brunswick). John Sim, a former NHL player with the Dallas Stars and New York Islanders, is the Director of Operations in Nova Scotia. They provide programming for both male and female hockey players and include speciality programs such as power skating, goaltending, and high-performance player development [22].

The participant sample (n=8) was comprised of minor hockey-level coaches who coach with the AHG and instruct youth aged 4-18 years. The participants were positioned in a head or leading coach role and had main responsibilities to coordinate team rosters, select assistant coaches, schedule practices, book ice times, and coordinate communications with both players and parents/guardians. Within the AHG organization, approximately 95% of head coaches instruct the co-ed and/or male teams, and approximately 75% of male coaches instruct the female teams [23].

This research used purposive sampling, as ice hockey coaches, both male and female, were invited to participate. Recruitment also involved notices through social media and direct email within the AHG organization. This qualitative research study utilized open-ended, semi-structured interview questions to collect data specifically regarding SEL training experiences of hockey coaches. There were small inducements of a coffee card available for participation. Participation was voluntary, and coaches were invited to complete an interview with a member from the research team. The interviews took place in person and/or online via Zoom, depending upon location and availability of each participant. The in-person interviews were conducted primarily within public places. Within the introduction segment of the interview, the researchers read a scripted informed consent, project purpose, and ethics brief from Cape Breton University. This informed the participants of the process and contacts for the CBU research ethics department should they have any questions. Informed consent was obtained prior to the start of each interview. The researchers also explained to participants that they were free to stop the interview at any time should they feel uncomfortable.

Participants were also invited to share their age and number of years coaching. Interview questions were not provided to participants in advance.

Regarding the participant demographic characteristics, 8 (100%) were male participants, and the age range was between 23 and 48 years. The average years of coaching was 9.6 years.

The qualitative interviews were comprised of 5 questions, which were developed to understand the SEL leadership training experiences of hockey coaches in Atlantic Canada.

Interview questions included the following:

1. Can you describe your position with the organization and what are your roles and responsibilities?
2. Can you briefly describe your ice hockey coaching training background? When and what types of training, the length and the content included?
3. 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?
4. In your coaching career, how often are you provided with leadership training? What types of training did you receive as Professional Development?
5. 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?

Field notes were taken by the researchers at the end of sessions to ensure key messages were highlighted, and sessions were also recorded with permission for transcription purposes. Ethics approval was obtained from Cape Breton University. Within the ethics approval, all data was collected and stored within the researchers' personal computers. Raw data was not accessible for public view.

3 Findings

Two main themes emerged from the analysis. Data were organized and analyzed using codes and a thematic approach. The emerging themes were larger and abstract, while the codes were one-word adjectives. Participants referred to a broad range of experiences to describe SEL and coaching training. Inductive content analysis was employed as this project included non-complex research, and the sample size (n=8) was small [24]. As noted by Williams and Moser [25], coding in qualitative research is comprised of processes that enable collected data to be assembled, categorized, and thematically sorted, providing an organized platform for the construction and development of meaning. Interviews were recorded (n=5) and notes were transcribed (n=3) for remaining interviews. The researchers reviewed the collected data which was shared with the team through MS Office Teams. The data analysis was based on word frequency and included verbs and adjectives. As an inductive process, the researchers became familiar with the data, generated initial codes, searched, and developed themes from the data and defined those main categories. Using an Excel spreadsheet, themes and codes were documented by the Primary Investigator. In this article, two main themes will be presented.

3.1 Theme 1: Training Requirements for Coaching Adolescents/Youth

Coaches reported (n=8) that they were head or lead coaches within the AHG organization. Ice hockey coaching training was received through a provincial/national entity. The analysis of data yielded a theme specific to importance of knowledge specific to adolescent behavior and youth development within a coaching context. 75% (6/8) of respondents noted they did not recall receiving training relating to this topic but would find this to be beneficial. Most of the training received was technical based, such as hockey drills and skills relating to shooting or positions. As Participant #4 explained:

> sometimes it can get it can get overwhelming when you're dealing with 17, 14, 15-year-olds that are just going through puberty, just starting to hit that. And then you have other kids that are already there and have hit it. The hormones are just flying.

Participant #2 stated that he utilized youth player development and positive psychology-based training, and it was self-initiated through online research web-based applications. Results also revealed that online training programming was limited or outdated and didn't match the current landscape of hockey players and coaches. Participant #1 explained that "the training is really old, and it really wasn't exciting... it's just online modules and... some old videos, but a lot of the materials are really, really old and really not interesting for anybody that's going to participate".

Table 1 describes codes and themes which were identified during the analysis. Coach training adolescents and youth was one theme which was evident. The need for current and relevant training to meet the needs of coaches who are coaching youth with diverse abilities and needs. 100% of coaches reported they did not receive specific SEL training. It was noted, however, that coaches (n=6) described receiving information about respect in hockey topics, diversity, and inclusive practice. Including SEL within a training and development practice is beneficial for children and youth to understand personal and social responsibility [26]. Engaging hockey coaches who spend countless hours coaching and mentoring children and youth with SEL skills could be a powerful addition to any hockey coaching training regime.

3.2 Theme 2: Effective Communication Practices and Processes

Communication was discussed by 100% of respondents during the interviews. Participants described the impact of coaching children and the importance of

Table 1: Coding and emerging themes

Codes	Emerging Themes
Course Content, Behaviour, Scheduling, Stress Management, Team Building, Children, Delivery of Training, Barriers, Format of Training	Training Coaching Adolescents and Youth
Relationships, Expectations, Comprehension, Listening, Inquiring, Providing Information, Happy, Encouraging, Discussions, Meetings	Effective Communication Practices and Processes

relationship building and expectations as it related to both coaching youth, but also communicating with parents.

Participant #5 explained: You know, they're getting into this teenage age . . . they don't want to listen to adults. They do, but they don't. And they always respect the coaches and love to promote. But finding that that that way to bridge that gap between these preteens that don't want to listen to adults".

Additionally, another response clearly outlined: "Some of the things they say they might make you might make you mad or might just annoy you to shreds" (Participant #4).

Addressing specific communication needs was also highlighted (n=3). Question #5 asked participants to discuss any additional or specific training. Understanding appropriate communication responses for players and parents is important. As an example, the following excerpt defines concerns from a coach: "She [child] did not play center, which then came to I'll call it almost came to blows with the parents getting in my face".

Communicating with guardians or parents is valuable and important. As described by Participant #6:

> setting expectations with the parents.. would be the parents, relationship, you get all, the all the issues with complaints and the blame so that's a huge part of minor hockey, to develop and trying to maintain that relationship... You know having parent meetings up front ongoing discussions with parents through the years and that gets a little bit easier the more you've coached.

The various types of communication were expressed during the interviews. Results indicated coaches utilized team meetings, individual and group emails, social media, and "Coffee with the Coach" panels for a general Q/A period. Open communication was noted as being extremely important to build trust but can also be a challenge without any professional development training. As Participant #7 explained:

> I have to talk to him [player] more than the others, but generally that's probably the biggest challenge how to how to chat and make those relationships without focusing on the game all the time.

The method in which coaches and players exhibit communication can be both verbal and non-verbal. Outbursts of aggression or demonstration of discontent can often result in challenges for the team. Coaches in this study acknowledged that derailing game plans or miscommunication can be difficult at times. Actions and reactions to situations on the ice may impact the team.

In summary, effective practices and processes for communication are an integral part of the coaching role. Guidelines and expectations contribute to this process. The following response identifies this challenge for coaches:

> we have to be prepared ...and parents sometimes are causing a barrier with playing and there's a lot of emotions and you really have to be honest with the kids you know be encouraging and how do you relate to the parents and how do you diffuse a situation... sometimes I'm getting a text in the middle of the night saying my son needs you and you know I'm the coach and trying to understand what does this mean?

4 Discussion

This study highlights the professional development and leadership training experiences of minor ice hockey coaches. Participants were members of the Atlantic Hockey Group organization in Atlantic Canada. Components of Social Emotional Learning (SEL) aspects were identified and explored. The data revealed that hockey training with aspects of SEL were limited and not specific or explicit to Social Emotional Learning. Unlike practitioners in many other fields of teaching or education, coaches in youth sport environments often have limited formal training or financial compensation for their work [27].

Within the analysis, training content for ice hockey coaches was described as primarily focused upon health and physical safety. Concussion protocol, dressing room safety, and physical environment of dressing room areas were reported. Hockey Canada and the Respect in Sport program was noted as an inclusive or diversity training program offered through coach training [28]. When asked about content of this module, participants could not recall specific modules that addressed diversity, inclusion, or communication. Participants did not recall receiving professional development (PD) training relating to SEL practices. Professional development can be defined as:

> Gaining new skills through continuing education and career training after entering the workforce. It can include taking classes or workshops, attending professional or industry conferences, or earning a certificate to expand your knowledge in your chosen field. [29]

As well, PD training was self-initiated and occurred in both online training and in-person formats. In terms of process, the overall findings of this study suggest that additional PD in areas relating to Social Emotional Learning would be beneficial, specifically with aspects of communication, diversity, inclusion, relationship building, and adolescent development. It has been argued that sport organizations do not seem to have systematic programs to teach life and sport skills [30, 31]. This study emphasized the important connections and impact within SEL, sport, and youth development. SEL has become a staple of positive youth development approaches within many educational settings [32]. SEL programming and awareness prepares students to move successfully through life transitions and is an equitable approach to supporting students of diverse backgrounds, including those of minority race or from disadvantaged families [33].

The findings of this study emphasize the importance of current professional development training required for the changing landscape of ice hockey in Canada. Providing youth hockey coaches with specific training relating to SEL may have the power to impact personal growth and leadership. Fostering effective relationships and mentoring youth may be impactful on coaches' abilities to bring additional success to the ice.

This study may align within hockey analytics and coaches' ability to motivate players to become successful leaders both on and off the ice. Building confidence through social skill development, improved communication practice, and leadership may equate in a hockey team's ability to improve upon their drills, scoring, and game planning.

Communication, adolescent development, parent and player expectations, and inclusion were focal points addressed by participants. Our study reported that coaches received technical training in ice hockey skill development, however limited or no training within aspects relating to SEL. Supplementing ice hockey training with PD sessions focused on SEL can not only educate hockey coaches about social skill development, but it can empower them to meet the diverse needs of youth and children whom they support. It is essential to foster relationship skills, self-management, and responsible decision-making in that it will engage managing emotions and exploring possible solutions through equity-focused conversations [13].

Another suggestion for future research involves parent communication and referees. Previous reports have suggested that these aspects can be beneficial in sport coaching [26].

Expanding on this study, the role of parents/guardians in relation to respecting communication with referees was often noted by researchers. Additionally, a future recommendation is for a study that explores the development for coaching various ability levels of players. Team building, social development skills, and using sport to build positive social relationships could be beneficial. Coaches also witnessed the pressure that some players experienced from parents to excel in the sport of hockey and the challenges that are associated with this process. Maintaining an atmosphere where youth could enjoy hockey and enhance social skill development was also highlighted by coaches as being an important area for development.

Limitations noted within the study included sampling of one hockey organization. All participants had coached hockey for the AHG in Atlantic Canada.

In summary, key findings from this study revealed that participant coaches were head or lead coaches within the AHG organization, received ice hockey coaching training from provincial minor hockey organizations, training format delivery included both online and in-person, explicit training specific to SEL was limited, effective communication was stated by all coaches as being an important aspect relating to coaching, and many coaches suggested the need for additional training within SEL, diversity, and inclusion.

5 Conclusion

The current study explored leadership and professional development training experiences of hockey coaches in Atlantic Canada. Results suggest that the majority of coaching training is received through both national and provincial associations. This training is specific to technical drills and skills for hockey development. Limited social skills or leadership training was received. Requirements and additional training that focus on inclusion, diversity, adolescent development, and communication within a hockey realm are noted. Using the tenets and foundations of SEL may provide an important benefit for professional development and leadership training for hockey coaches in Atlantic Canada.

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

References

1. Cairnie, J.: Truth and reconciliation in postcolonial hockey masculinities. Can. Lit. 237, 103–119 (2019).

2. International Ice Hockey Federation. Survey of players (2018). http://webarchive.iihf.com/iihf-home/the-iihf/survey-of-players/index.html

3. Adams, C., Leavitt, S.: "It's just girls' hockey": troubling progress narratives in girls' and women's sport. Int. Rev. Sociol. Sport. 53, 152–172 (2018). https://doi.org/10.1177/1012690216649207

4. Hockey Canada. Statistics & history (2022). http://www.hockeycanada.ca/en-ca/Hockey-Programs/Female/Statistics-History.aspx

5. 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. Int. Sport Coaching J. 7, 335–346 (2020). https://doi.org/10.1123/iscj.2019-0096

6. Lara-Bercial, S., Mallett, C.J.: The practices and developmental pathways of professional and Olympic serial winning coaches. Int. Sport Coaching J. 3, 221–239 (2016). https://doi.org/10.1123/iscj.2016-0083

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

8. Gould., D. Martin, E., Walker, L.: A season long investigation of social emotional learning associated with high school basketball participation. J. Appl. Sport Psychol. 34, 1102–1124 (2022). https://doi.org/10.1080/10413200.2021.1955421

9. Zins, J.E., Bloodworth, M.R., Weissberg, R.P., Walberg, H.J.: The scientific base linking social and emotional learning to school success. J. Educ. Psychol. Consult. 17, 191–210 (2007). https://doi.org/10.1080/10474410701413145

10. Weissberg, R.P., Durlak, J.A., Domitrovich, C.E., Gullotta, T.P.: Social and emotional learning: past, present, and future. In: Durlak, J.A., Domitrovich, C.E., Weissberg, R.P., Gullotta, T.P. (eds.) Handbook of Social and Emotional Learning: Research and Practice, pp. 3–19. Guilford Press, New York (2015)

11. Domitrovich, C.E., Durlak, J.A., Staley, K.C., Weissberg, R.P.: Social-emotional competence: an essential factor for promoting positive adjustment and reducing risk in school children. Child Dev. 88, 408–416 (2017). https://doi.org/10.1111/cdev.12739

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

13. Collaborative for Academic, Social, and Emotional Learning (CASEL). CASEL's SEL Framework: What Are the Core Competence Areas and Where Are They Promoted? CASEL (2020)

14. Anderson-Butcher, D. et al.: Social-emotional learning interventions in youth sport: what matters in design? Child Adolesc. Social Work J. 38, 367–379 (2021)

15. Collaborative for Academic, Social, and Emotional Learning (CASEL). CASEL guide: effective social and emotional learning programs (2015).

16. Camiré, M.: Benefits, pressures, and challenges of leadership and captaincy in the National Hockey League. J. Clin. Sport Psychol. 10, 118–136 (2016)

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

18. Bronfenbrenner, U.: Developmental ecology through space and time: a future perspective. In: Moen, P., Elder, Jr., G.H., Luscher, K. (eds.) Examining Lives in Context: Perspectives on the Ecology of Human Development, American Psychological Association, Washington, DC (1995)

19. Darling-Hammond, L., Cook-Harvey, C.M.: Educating the Whole Child: Improving School Climate to Support Student Success, Learning Policy Institute, Palo Alto, CA (2018)

20. Luna, P., Guerrero, J., Cejudo, J.: Improving adolescents' subjective well-being, trait emotional intelligence and social anxiety through a programme based on the sport education model. Int. J. Environ. Res. Public Health. 16, 1821–(2019)

21. Jones, S.M., 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. National Commission on Social, Emotional, and Academic Development, The Aspen Institute (2017)

22. Atlantic Hockey Group. About the Atlantic hockey group: our history (2023). https://www.atlantichockeygroup.com/

23. Bourgeois, B.: Personal email, May 24, 2023

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

25. Williams, M., Mosher, J.: The art of coding and thematic exploration in qualitative research (2019). http://www.imrjournal.org/uploads/1/4/2/8/14286482/imr-v15n1art4.pdf

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

27. Schlechter C.R., Rosenkranz, R.R., Milliken, G.A. et al.: Physical activity levels during youth sport practice: does coach training or experience have an influence? J Sports Sci. 35, 22–28 (2017).

28. Hockey Canada. Respect in sport (2022). https://www.hockeycanada.ca/en-ca/hockey-programs/coaching/essentials/faq/respect-in-sport

29. Parsons, L.: Why is professional development important? (2022). https://professional.dce.harvard.edu/blog/why-is-professional-development-important/

30. Camiré, M., Trudel, P., Bernard, D.: A case study of a high school sport program designed to teach athletes life skills and values. Sport Psychol. 27, 188–200 (2013). https://doi.org/10.1123/tsp.27.2.188

31. Petitpas, A.J., Cornelius, A.E., Van Raalte, J.L. et al.: A framework for planning youth sport programs that foster psychosocial development. Sport Psychol. 19, 63–80 (2005)

32. Elias, M.J., Kranzler, A., Parker, S.J., Kash, V.M., Weissberg, R.P.: The complementary perspectives of social and emotional learning, moral education, and character education. In: Nucci, L., Krettenauer, T., Narvaez, D. (eds.) Handbook of Moral and Character Education, 2nd ed., Routledge, New York (2014)

33. Taylor, R.D., Oberle, E., Durlak, J.A., Weissberg, R.P.: Promoting positive youth development through school-based social and emotional learning interventions: a meta-analysis of follow-up effects. Child Dev. 88, 1156–1171 (2017)

Analyzing Passing Metrics in Ice Hockey using Puck and Player Tracking Data

David Radke, Jaxin Lu, Jackson Woloschuk, Tin Le[†],
Daniel Radke, Charlie Liu, and Tim Brecht

Cheriton School of Computer Science, University of Waterloo
University of Calgary[†]

Abstract. Traditional ice hockey statistics are inherently biased towards offensive events like goals, assists, and shots. However, successful teams in ice hockey require players with skills that may not be captured using traditional measures of performance. The adoption of puck and player tracking systems in the National Hockey League (NHL) has significantly increased the scope of possible metrics that can be obtained. In this paper, we compute recently proposed passing metrics from 1221 NHL games from the 2021-2022 season. We analyze the distributions of values obtained for each player for each metric to understand the variance between, and within, different positions. We find that forwards tend to complete fewer passes with smaller passing lanes, while defensemen pass to forwards significantly more than their defensive partners . Additionally, because these new metrics do not correlate well with traditional metrics (e.g., assists), we believe that they capture aspects of players' abilities that may not appear on the game sheet.

1 Introduction

The idea of using quantitative evidence to understand player tendencies and performance to inform management and strategic decisions has existed in sports for several decades [9]. In sports classified as "striking games", such as baseball, analytics has transformed team operations and strategies [4]. This influence has lagged behind in "invasion games" such as football (soccer), basketball, handball, and ice hockey due to limitations in data collection and the complexities of the sport. Traditional (publicly available) statistics captured in ice hockey revolve around easily measurable offensive events (i.e., goals or shots) leading to the performance of offensive players being disproportionately captured. Successful teams in ice hockey, like all invasion games, require players with diverse abilities that existing offensively biased metrics do not capture, such as passing. This limited information makes constructing teams using quantitative evidence more difficult. The recent implementation of the puck and player tracking (PPT) systems in the National Hockey League (NHL) has led to several new metrics to quantify player behavior [12, 13]. In this paper, we utilize a larger dataset to study how passing metrics can be utilized to understand the variance in behavior among players and players at different positions (metrics with larger variance may provide more opportunities to find under-valued players).

The main motivation behind the development of passing metrics in ice hockey was to capture other player contributions that might not show up on a game sheet [12]. Understanding how players compare to each other within the distribution of passing metrics provides valuable context for team building and management. We perform a deeper analysis into recently proposed ice hockey metrics from NHL puck and player tracking (PPT) data to show how passing metrics can be used to identify diverse behaviors among individuals. The contributions of this work are:

- We perform significant amounts of data cleaning to calculate passing metrics using PPT data from 1221 games in the 2021-2022 NHL season.
- We analyze the distributions of various passing metrics for forwards and defensemen. This provides insights into how much better highly-ranked players are when compared with other players.
- We find that after normalizing for ice time, forwards tend to complete fewer passes than defensemen and have smaller passing lanes, whereas defensemen complete significantly more passes to forwards and overtake more opponents.
- We show that the number of players overtaken with completed passes and the size of the passing lanes for completed passes do not correlate well with traditional offensive-oriented statistics like assists. We believe this demonstrates that some of the our metrics capture aspects of players' abilities that might not show up on the game sheet.

2 Related Work

Understanding how multiple players/agents work together most effectively is a significant area of research in organizational psychology and AI [1, 15]. A general finding is that group diversity, role specialization, and cohesiveness is important for group performance [7, 1, 23, 14]. Similar results have been found in football and sports analytics. Those analytics have focused on the performance of groups of players together [8, 11, 3, 10]. We use football to refer to *association* football (also known as soccer), not American football.

The implementation of passing metrics in football allows the analysis of a player's decision-making and passing ability [20, 19], ability to overtake players with passes [21], impact on scoring probability [5], and ability to act under pressure [2]. In a low-scoring game like football, these models provide insight into players' behaviors independently from offense and enables team building with diverse skills. Similar advancements in ice hockey have analyzed passing lane probabilities [17], as well as passing scenarios and pressure [12, 13].

Despite the development of models that use PPT data in ice hockey, no previous work has analyzed passing models to help understand general distributions, trends, or differences among players in the NHL. In this paper, we calculate various passing metrics from recent work [12, 13] for 1221 games of the 2021-2022 NHL season. We analyze the distributions of each metric among players at different positions and within each position. Furthermore, we cross-reference related metrics to gain insight into how individual players behave with respect to multiple metrics.

3 Puck and Player Tracking Data

In the NHL, hockey is played on an ice surface that is 200 feet long and 85 feet wide. Tracking data is collected by SportsMEDIA Technology [18] (a partner of the NHL). They then derive event-level data (including completed passes) from the location tracking data. These event labels contain information about the time of the event and the identities and locations of the players involved. This paper focuses specifically on completed pass events. We have been granted early access by the NHL to the first full season for which the NHL used the PPT system (2021-2022). The PPT data and our resulting metrics are considered unofficial by the NHL, as the models used for creating event labels continue to be validated and improved. Additionally, the process of making statistical data official requires approval in the collective bargaining agreement, an ongoing process that has not been completed at this time. As a result, we do not provide information about individual player's metrics. Also note that this data may differ from other datasets that contain complete and/or incomplete passes (e.g., a hand labeled dataset). We have processed 1221 of the 1312 regular season NHL games.[1]

Location data is collected through tracking technology that is embedded into pucks and inserted into the sweaters of each player (on the back of the sweater, slightly right of the center of the shoulders). Location information contains x, y, and z-coordinates to record locations in 3-dimensional space. The x and y locations are relative to center ice (which is 0, 0). It is our understanding that when tested, the margin of error for the x and y coordinates is about 3 inches (the diameter of the puck) and very often as little as 1 inch. This is accurate enough for our purposes, as a puck traveling at speeds between 30 and 100 MPH would travel between 8 inches and 28 inches, respectively between readings. Further, our metrics are not overly sensitive to small changes in the puck's location. The z coordinates (not used in this paper) are relative to the surface of the ice. Location data is recorded 60 times per-second for the puck and 12 times per-second for each player on the ice, resulting in a total of about 734,400 location readings of interest in a 60-minute game. Additionally, location data is obtained once per-second for players that are determined to be off of the ice. We interpolate all puck and player locations to 100 readings per-second to more easily identify the positions of all on-ice entities at precisely the same time.

[1] Some games could not be processed due to issues with the data sets and/or some special cases that our code hasn't yet handled.

4 Background of Metrics

We briefly discuss the passing models used to derive the metrics in this paper. We refer the reader to the original paper [12] for further passing model details and to [13] for extensions and improvements to the original passing lane model. To ensure comparisons of different players are fair and are not simply a measure of ice time, we normalize our metrics by time on ice and/or games played, where appropriate (e.g., as is done in Section 6).

4.1 Passing Lane Model

The passing lane model we use in this paper is originally proposed in [12] and enhanced in [13]. The model uses the spatial locations of players in PPT data to estimate the available space between a passer p and any receiver r.

Figure 1 (adapted from [12]) shows the passing lane shape for a direct pass from p to r with three opponents. For each passing event, the model constructs a teardrop-like passing lane shape around the passer p and extending beyond the location of the receiver r (shaded regions). The size of the passing lane is determined by the nearest opponent to the pass and assigns a non-negative real-numbered value γ to be the openness of the pass. Figure 1 shows three passing lanes with respect to each of the three opponents. The γ value of this pass is $\gamma = 0.6$, since o_1 restricts the passing lane the most. We use the enhanced version from [13], where the *expected* locations of the receiver r and all opposing players based on current velocities are used to determine the passing lane. The enhancement also considers indirect passes off the boards. We developed a new constant-time algorithm to directly calculate γ instead of the previous binary search method. Refer to our previous paper [13] for more details.

4.2 Pass Overtaking Model

Previous work proposed and implemented models to understand progressing the puck beyond opponent players with passes [12]. At a high level, this model is represented as a zero-sum game, where a passer p gains a positive value for overtaking opponents with passes, and each opponent overtaken o receives a corresponding negative value. Formally, for a completed pass from p to r, if $\delta(x, y)$ is Euclidean distance between location x and y and $\texttt{NET}$ is the center of the entrance to o's net, o is considered overtaken if $\delta(p, \texttt{NET}) > \delta(o, \texttt{NET})$ and $\delta(o, \texttt{NET}) > \delta(r, \texttt{NET})$.[2] Because defensemen have greater opportunities to overtake more players, the model uses the fraction of players that are possible to be overtaken as the allocated credit. For example, if there are 3 players between p and the net (not counting the goalie) and the pass overtakes 2 opponents, the pass overtake value is 0.67. The passer p receives a positive value of $+0.67$ and each of the two overtaken players receives a negative value of -0.33 while the remaining non-overtaken player is unchanged.

[2] See Section 8 for variations we plan to consider in future work.

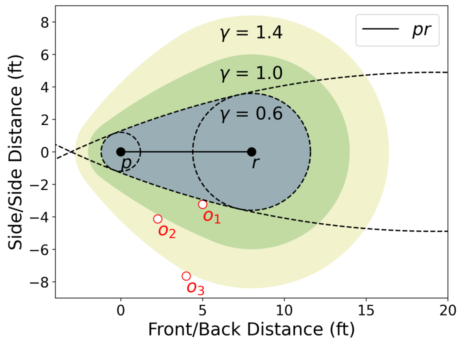


Fig. 1: Passing Lane.

Fig. 2: Adapted from [12]. The passing lane model for direct passes. The passing lane (shaded regions) surrounds the passer p and receiver r. The size and shape of this lane scales to the nearest opponent o (we show three examples of passing lanes with respect to three opponents). We use an expanded version that incorporates expected movement and indirect passes [13].

These values are aggregated into various metrics, including OVT (overtake total), BTT (beaten total), and PPM (passing plus-minus), calculated as PPM = OVT - BTT. We also calculate OVA, the average fraction of players overtaken with each pass ($\frac{\text{OVT}}{\text{num\_passes}}$). Because there can be significant differences in the number of games played by different players, we use average values per game where appropriate. This ensures a fair comparison when examining and comparing different players.

5 Data Cleaning

When beginning our analysis we found several anomalies that needed to be corrected. Specifically, when using the timestamps associated with a fair number of completed passes, the puck was located at a relatively large distance from the passing player (e.g., significantly outside the reach of the player). To mitigate this issue, we performed a pass timestamp correction phase to better identify and adjust the time at which the event occurred. Adjusting these timestamps is also important to correctly identify the locations of all players on the ice at the time of the event. This is critical to obtain accurate passing metrics. All results in this paper are computed after adjusting the timestamps, which has significantly improved our metrics.

Our adjustment process begins by finding the timestamp for an event t in the PPT data. At a high level, our approach is to find a more accurate timestamp t' where the puck is sufficiently close to the passing player (i.e., within reach of the player).We determined a threshold of $\delta(p, puck) \leq 4$ feet to be a reasonable value, based on discussions with people at the NHL and personal measurements.

Metric	Description
avgOVT_20	The sum of the fraction of opponents overtaken by a player's passes. We scale to 20 minutes of ice time and average per-game.
avgBTT_20	The sum of the fraction a player was overtaken by opponents' passes We scale to 20 minutes of ice time and average per-game.
avgOVA	The average fraction of opponents overtaken by a pass in a game. We average this value per-game.
avgPAA	Average γ (passing lane) value for completed passes. We average this value per-game.

Table 1: Summary of passing metrics discussed in this paper. Additive metrics (totals; end with "T") are averaged over players' games played ("avg") and scaled to 20 minutes per-game if necessary ("_20").

Any passes that could not be corrected using this technique are omitted from our dataset. This was only about 2.6% of the total number of completed passes. There are several possible ways to improve the accuracy of this approach including examining changes in the direction and speed of the puck. However, determining the accuracy of various techniques requires knowing ground truth, as a result this is a topic for future research.

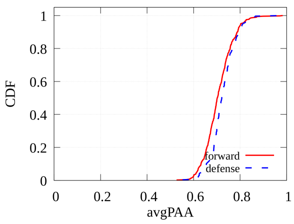
6 Distribution Analysis of Passing Metrics

In the original work where we proposed these passing models we only had access to smaller PPT datasets so we did not conduct a detailed analysis for players [12]. In this paper, we analyze 1221 games and examine whether or not there are differences in passing metrics between forwards and defensemen and study the differences among individual players within the same position. We provide a summary of the metrics we analyze in Table 1. Our dataset includes 1000 players. To ensure that we have a sufficient sample size for various metrics we exclude players that did not play in at least 10 games and average at least 10 minutes of ice time per game. This reduced our dataset to 750 players (478 forwards and 272 defensemen). Because our work in this paper focuses on passing , we do not include goaltenders in any of our metrics or player counts.

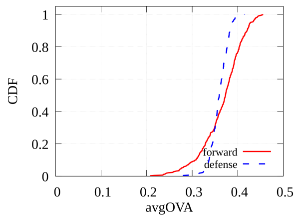
To allow for fair comparisons among players that receive different amounts of ice time (since some metrics correlate with ice time) we normalize metrics (where appropriate) to 20 minutes per game. For each of the metrics in Table 1 we average over a player's games. Thus, a metric such as OVT, the total fraction of opponents a player overtakes with their passes, will be represented as $avgOVT\_20$: averaged over a player's games ("avg") and scaled to 20 minutes of ice time per-game ("_20"), where appropriate.

6.1 Distributions of Metrics Based on Position

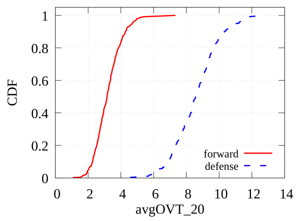
We perform the Welch t-test [22] to analyze how the distributions of metrics vary between forwards and defensemen. When rejecting the null hypothesis for the mean values of a metric being equal between forwards and defensemen at a p-value of 0.05, we find that the mean between forwards and defense for both traditional statistics (e.g., goals, assists, points, shots, and shots blocked) and the new passing metrics in Table 1 are sufficiently different for every metric except for hits (which we do not consider in this paper). As a result, we analyze forwards and defensemen separately and use cumulative distribution functions (CDFs) to analyze the variance of distributions at each position.



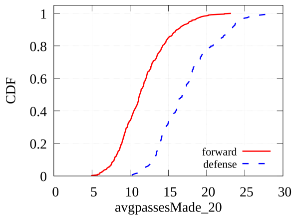
(a) avgPAA.



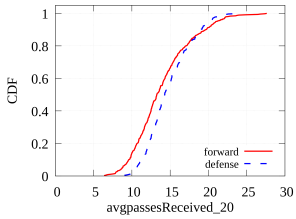
(b) avgOVA.



(c) avgOVT_20.



(d) avgPassesMade_20.



(e) avgPassesReceived_20.

Fig. 3: CDFs plots for passing metrics separated by position, including (a) avgPAA: the average γ value for completed passes (lower indicates smaller passing lanes). (b) avgOVA: the average fraction of opponents overtaken by a pass (larger is better). (c) avgOVT_20: the total fraction of opponents overtaken by a pass (larger is better). Metrics for each player are averaged over the number of games they have played ("avg") and when appropriate scaled to 20 minutes of playing time per-game ("_20"). (d) The average passes made by players per 20 minutes. (e) The average passes received by players per 20 minutes.

Passing Metrics Figure 3a shows the CDF for avgPAA, the per-game average γ value (passing lane size) for completed passes. Our results show that forwards and defensemen have distributions with similar shapes; however, the median defensemen tends to complete passes with slightly larger passing lanes. The forwards with the lowest avgPAA complete passes with about 47% smaller passing lanes than the forwards with the highest avgPAA.

Figure 3b shows the CDF for avgOVA, the average fraction of opponents a player overtakes per-pass, per-game. Higher values of avgOVA suggest the player overtakes a higher fraction of opponents with each pass (i.e., a stretch pass beating four of five players gives $\frac{4}{5} = 0.8$, while beating only the last defender gives $\frac{1}{1} = 1$). The 30th percentile values of each position are similar (about 0.35). Defensemen have lower variance in avgOVA with the range from the 20th percentile to 80th percentile being from 0.34 to 0.37 per-pass (34% to 37%) of the possible players per-pass. Comparatively, forwards have over double the variance than defensemen in avgOVA and the forwards with the highest avgOVA have over double the overtake value per-pass compared to the lowest forwards (0.45 compared to 0.21). The 20th percentile of forwards overtake an average of 33% of the possible players per-pass and the 80th percentile forwards overtake and average of about 40% of the possible players per-pass. The larger variance among forwards is likely caused by forwards typically having fewer opponents to overtake (2 or 3) compared to defensemen (4 or 5). We note that lowest percentile forwards are players that tend to make fewer than five passes per 20 minutes. We acknowledge that there may exist some players within our dataset that circumvent the intent of our filter and if they have a low number of passes, that could skew the distributions of some metrics. Future work could consider filtering techniques to remove players with too few passes.

Figure 3c shows the CDF for the avgOVT_20, the per-game average of the *total* fraction of opponents overtaken with passes normalized for 20 minutes of ice time. Higher values of avgOVT_20 suggest the player overtook a large fraction of opponents with their passes throughout a game. Our results show that the median defenseman achieves 2.8 times higher avgOVT_20 than the median forward (comparing 3 for forwards to 8.5 for defensemen). This difference of 5.5 avgOVT_20 increases to about 6 at the 80th percentile of forwards and defensemen (comparing 3.9 for forwards to 9.5 for defensemen). This change in the differences means the top defensemen for avgOVT_20 overtake more opponents compared to other defensemen than the top forwards compared to the rest of the forward population. Since the median avgOVA value for forwards is 4% higher than the median defensemen (see Figure 3b), we can conclude that higher values of avgOVT_20 for defensemen indicate that they complete more passes than forwards. This is confirmed in Figure 3d which shows the distribution of completed passes per 20 minutes. The shapes of the distributions among each position are almost identical, but defensemen tend to complete about five more passes than forwards at every percentile. The median forward completes about 12 passes per 20 minutes, whereas the median defenseman completes about 17. The forwards that complete the most passes complete up to 23 passes per 20 minutes 92% more than the forward median) and the defensemen that complete the most passes complete up to 27 passes (59% more than the defense median).

On average, despite defensemen completing roughly five more passes each than forwards, both positions tend to receive about the same number of passes

(Figure 3e). Comparing Figures 3d and 3e allows us to draw an interesting conclusion: defensemen complete passes to forwards significantly more often than to their defensive partner.

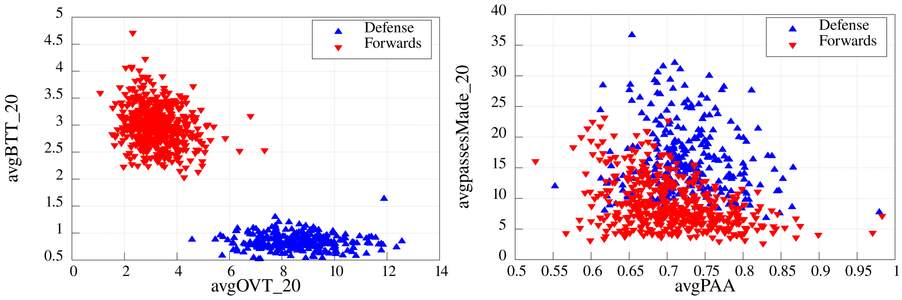
To understand this insight, consider that at even strength, the players that a forward can pass to are the two other forwards and the two defensemen. Assuming each of the other players is equally likely to be chosen (which may not be true), the probability of passing to a forward or defensemen is equal at 0.5. However, for defensemen there are three forwards and one defensemen to choose from. Again assuming the probability of passing to each of the other four players is equally likely (i.e., 0.25), the probability of passing to a forward is 0.75 and their defensive partner is 0.25. For the average pass reception curves for defensemen and forwards to be similar (as seen in Figure 3e) defensemen must complete passes to forwards three times more often. Considering that defensemen typically complete about five *more* passes than forwards (Figures 3d), defensemen must pass to forwards even more. Since these passes are likely up-ice, the higher frequency of passes from defensemen to forwards must be the main reason for high values of avgOVT among defensemen (Figures 3c).

Takeaways: We find that forwards make passes with slightly smaller passing lanes than defensemen. The variance among forwards for overtaking opponents with a pass (avgOVA) is significantly larger than with defensemen; however, the median forwards are only 4% higher than the median defensemen in avgOVA. Despite slightly lower median avgOVA, defensemen accumulate significantly higher *totals* for overtaking opponents (avgOVT_20) and complete about 5 more passes each game compared to forwards. Using Figures 3d and 3e, we find that defensemen pass to forwards significantly more than to their defensive partners.

6.2 Analyzing Player Differences

In this section we analyze individual players across a variety of metrics to gain insights into differences among players. One of the main passing metrics derived in our previous work [12] and discussed in Section 4 is passing plus-minus (PPM), defined as PPM = OVT - BTT. PPM gives insight into if a player overtakes more opponents than they are overtaken themselves; however, the metric removes additional context that may be important when understanding player behaviors. For example, a player that rarely overtakes opponents while also never being overtaken could have the same PPM value as a player that overtakes many opponents but often gets overtaken.

Figure 4a compares the two components of PPM to analyze the distribution of players along these two dimensions. The x-axis shows avgOVT_20, the total fraction of opponents that a player overtakes with their passes (per-game average, higher is better) and the y-axis shows avgBTT_20, their total fraction of being overtaken by opponents (per-game average, lower is better). Red triangles represent forwards and blue triangles represent defensemen. Players in the lower right corner overtake more opponents while not being overtaken by many opposing team passes. Analyzing where players are in these distributions may be important when constructing forward lines or defensive pairings as a coach, or a roster as a manager.



(a) avgOVT_20 vs. avgBTT_20.

(b) avgPAA vs. avgPassesMade_20.

Fig. 4: (a) The two components of passing plus-minus (PPM; presented in Section 4.2). The avgOVT_20 metric (overtake total; x-axis), the total (per-game average) fraction of opponents a player overtakes with their passes and avgBTT_20 (beaten total; y-axis), the total (per-game average) fraction that players are overtaken by opponents. (b) The average γ value (passing lane) for a player's completed passes (avgPAA; x-axis) compared to players' average completed passes per 20 minutes.

Figure 4a shows that there is diversity (or variation) among forwards with respect to both avgOVT_20 or avgBTT_20. Table 2 shows the avgOVT_20 and avgBTT_20 values with 95% confidence intervals for the forwards and defensemen with the highest, median, and lowest values for each metric. None of the confidence intervals for the three forwards intersect for either metric; thus, we can confirm that there exist forwards with differences that are statistically significant. In comparison, Figure 4a shows that defensemen mostly vary along the dimension of how they overtake opponents with passes (avgOVT_20). Table 2 confirms that the confidence intervals for defensemen do not intersect for avgOVT_20 but do intersect for avgBTT_20. Therefore, we conclude that defensemen mostly distinguish themselves from their peers by overtaking more opponents with their passes (avgOVT_20).

Figure 4b compares the per-game average value of γ (passing lane size) for a player's completed passes (avgPAA_20; x-axis) and the average number of passes made by that player (avgPassesMade_20; y-axis). For both forwards and defensemen, players that complete the most passes (higher on the y-axis) tend not to have the lowest or highest values of avgPAA (x-axis) compared to the other players within their position. This implies that the players who complete a large number of passes do so in situations that are not anomalous (i.e., they are not mostly passing in easier situations).

Takeaways: There exists diversity among forwards with respect to overtaking opponents with passes and being overtaken by opponent passes. Defense-

Player	avgOVT_20	95% CI	avgBTT_20	95% CI
Fwd. Highest	7.28	±1.84	4.66	±1.09
Fwd. Median	3.16	±0.42	3.00	±0.19
Fwd. Lowest	1.05	±0.43	2.05	±0.17
Def. Highest	12.46	±1.91	1.58	±0.73
Def. Median	8.57	±0.60	0.83	±0.24
Def. Lowest	4.55	±0.52	0.51	±0.10

Table 2: Analyzing the mean and 95% confidence intervals for the highest, median, and lowest values for forwards and defensemen for avgOVT_20 and avgBTT_20. Our results show diversity among forwards with respect to both metrics while the highest defensemen tend to mostly separate themselves from their peers with respect to avgOVT_20.

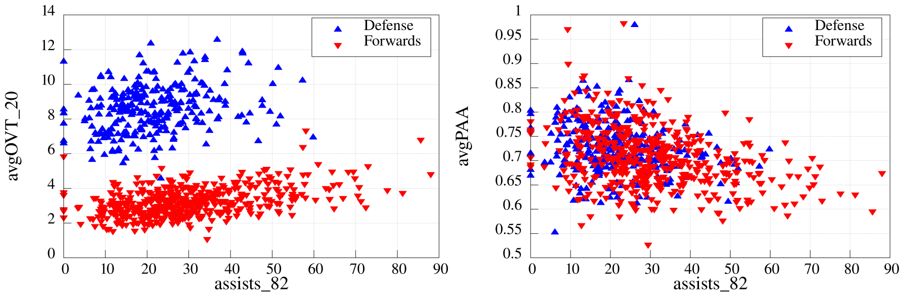
men mainly separate themselves from their peers by overtaking more opponents, while there is less distinction with how defensemen are overtaken by opponents. At both positions, players that complete the most passes tend to do so with an average passing lane size instead of completing a disproportionate amount of easier passes with bigger passing lanes.

7 Comparative Analysis

Inspired by the work on "Meta-Analytics" (to examine stability, discrimination and independence of metrics) proposed by Franks *et al.* [6], we present a simple analysis of some of our metrics to show that avgOVT_20 and avgPAA do not correlate well with assists (i.e., to provide some indication of independence from a traditional offensive oriented statistic). We also compare the avgOVA, avgOVT_20, and avgPAA metrics obtained from the first 50% of the games with the same metrics computed across the last 50% of the games we have processed (to examine the stability of those two metrics). We divide games using the unique value assigned to each game (game id) which are typically ordered by scheduled date. Note that because a small number of games were postponed due to COVID-19, the split may not be precisely by the date games were played. In the future we plan to conduct an in depth analysis of all of our metrics (and other existing statistics) using the "Meta-Analytics" framework.

7.1 Comparison with Traditional Statistics

Figure 5a compares assists_82 (normalized to 82 games with 20 minutes per-game) and avgOVT_20 (the sum of the fraction of opponents overtaken by a player's passes normalized to 20 minutes). Advancing the puck and overtaking opponents is a valuable aspect in invasion games like ice hockey [16]. Figure 5a shows there exists many players at both positions who overtake a significant number of opponents with completed passes who do not record a large number of assists. These players with high avgOVT_20 values may not always show up on a game sheet; however, they may be playing important roles on their team.



(a) assists_82 vs. avgOVT_20

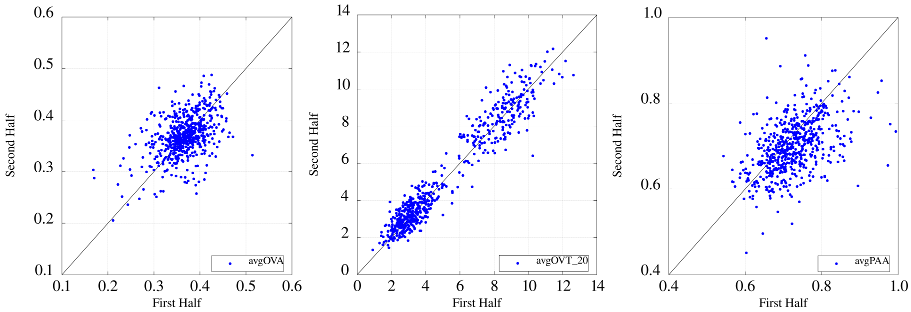
(b) assists_82 vs. avgPAA

Fig. 5: (a) Assists projected to 82 games and 20 minutes per-game (assists_82) versus the total fraction of opponents overtaken by passes (avgOVT_20; per-game average; higher is better). (b) Projected assists (assists_82) versus the average value of γ for completed passes (avgPAA; per-game average; lower means smaller lanes).

Figure 5b compares assists_82 with avgPAA for players' completed passes in a game (avgPAA is the average γ value, or passing lane size; lower indicates smaller lanes). Note that there are no players with both high avgPAA and high assists_82 (i.e., no players in the top right of Figure 5b). However, many players with the highest assists_82 values have relatively low avgPAA (between 0.59 and 0.70). This may suggest a connection between recording many assists and being able to complete passes with smaller lanes. In future work we plan to examine this question more closely by separating, studying and comparing passing lanes for completed passes that result in assists. Again, we believe that considering traditional offensively-oriented statistics for a player could reduce one's ability to see other potentially important skills.

7.2 Evaluating Stability

Figure 6 compares metrics computed over the first 610 games with the same metric computed over the last 611 games. If the metrics obtained for each player during the first half of the games were able to perfectly predict the metric computed over the second half of the games, all data points would fall exactly on the diagonal line. These graphs indicate that the avgOVA and avgPAA metrics are well correlated across the two halves of the season (their correlation coefficients, r, are 0.87 and 0.89, respectively). The avgOVT_20 metric is strongly correlated with $r = 0.99$. For comparison we found (details and graphs have been excluded for brevity) that the correlation coefficient for players' points is $r = 0.80$ and for goals is $r = 0.72$. This indicates that our new metrics are more stable (i.e., future values may be more predictable) than points and goals.



(a) avgOVA (b) avgOVT_20 (c) avgPAA

Fig. 6: Comparing different metrics from the first half of the games with the same metric computed over the second half of the games. Points on the diagonal line are perfectly correlated.

8 Discussion

While we perform an extensive analysis of several metrics and their distributions across players, our work has several limitations. One limitation is that we do not consider different factors such as coaching style (or team systems), manpower (e.g., even strength or not), goal differential, time of the game, and play location that may provide further insights. Future work may consider analyzing these scenarios separately.

Another limitation is the aggregation of metrics while including players with few samples. We filter our dataset by excluding players that don't receive a minimum average amount of ice time per game or have not played a minimum number of games. However, among the unfiltered players, some players recorded relatively few completed passes . Future work could apply additional filters (e.g., filtering players by a minimum number of samples).

Similar to the limitations with previous work [12, 13], we are only able to analyze completed passes. In the future we hope to discern or obtain information about unsuccessful passes. Additionally, our model for overtaking opponents does not consider potentially valuable passes such as those from close to (or behind) the net to the slot area, or east-to-west passes on odd-man-rushes, as overtaking opponents. Future work may adapt our model to include these types of passes.

9 Conclusions

Traditional ice hockey statistics disproportionately capture the offensive perspective of players. Understanding other characteristics of players' behaviors is important for constructing forward lines, defensive pairings, or entire teams. In this paper, we analyze several recently proposed passing metrics using PPT data from 1221 games of the NHL 2021-2022 season. We find that forwards tend to complete passes with slightly smaller passing lanes compared to defensemen; however, defensemen complete more passes and overtake more opponents. Examining players by comparing their scores on the basis of two metrics reveals the diversity of behavior among players with regards to pass overtaking and being overtaken by passes. Finally, because these new metrics do not correlate well with traditional metrics, we believe they capture aspects of players' abilities that may not appear on a traditional game sheet. This analysis may be of significant interest to coaches and managers as they attempt to construct successful teams.

Acknowledgments

We thank Rogers Communications for providing funding for this project. In addition, this research is partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), two University of Waterloo Math Undergraduate Research Awards and a University of Waterloo Undergraduate Research Fellowship. We thank the members of the National Hockey League's Information Technology Team and the Stats and Information Team for their participation in fruitful discussions and their insights related to this work. From AWS, we thank Candi Jeronimo for providing us with credits and Varad Ram for technical assistance that enabled us to process the games using AWS. We thank Neel Dayal and Jonah Eisen from Rogers Communications for making this project possible. We thank Robby Brecht for his feedback on an earlier draft of this paper. Finally, we thank the anonymous reviewers for their constructive and helpful feedback.

References

1. Andrejczuk, E., Rodriguez-Aguilar, J.A., Sierra, C.: A concise review on multiagent teams: contributions and research opportunities. Multi-Agent Systems and Agreement Technologies (2016)
2. Andrienko, G., Andrienko, N., Budziak, G., Dykes, J., Fuchs, G., Landesberger, T.V., Weber, H.: Visual analysis of pressure in football. Data Mining and Knowledge Discovery **31**, 1793–1839 (2017)
3. Bransen, L., Van Haaren, J.: Player chemistry: Striving for a perfectly balanced soccer team. arXiv preprint arXiv:2003.01712 (2020)
4. Elitzur, R.: Data analytics effects in major league baseball. Omega **90**, 102001 (2020)

5. Fernández, J., Bornn, L., Cervone, D.: Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer. In: 13th MIT Sloan Sports Analytics Conference (2019)
6. Franks, A., 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**(4), 151–165 (2016)
7. Hong, L., Page, S.E.: Groups of diverse problem solvers can outperform groups of high-ability problem solvers. Proceedings of the National Academy of Sciences **101**(46), 16385–16389 (2004)
8. Ingersoll, K., Malesky, E., Saiegh, S.M.: Heterogeneity and team performance: Evaluating the effect of cultural diversity in the world's top soccer league. Journal of Sports Analytics **3**(2), 67–92 (2017)
9. Lewis, M.: Moneyball: The art of winning an unfair game. WW Norton & Company (2004)
10. Ljung, D., Carlsson, N., Lambrix, P.: Player pairs valuation in ice hockey. In: MLSA@PKDD/ECML (2018)
11. Nsolo, E., Lambrix, P., Carlsson, N.: Player valuation in european football. In: Machine Learning and Data Mining for Sports Analytics: 5th International Workshop, MLSA 2018, Co-located with ECML/PKDD 2018, Dublin, Ireland, September 10, 2018, Proceedings 5. pp. 42–54. Springer (2019)
12. Radke, D.T., Radke, D.L., Brecht, T., Pawelczyk, A.: Passing and pressure metrics in ice hockey. Workshop of AI for Sports Analytics (2021)
13. Radke, D., Brecht, T., Radke, D.: Identifying completed pass types and improving passing lane models. In: Linköping Hockey Analytics Conference. pp. 71–86 (2022)
14. Radke, D., Larson, K., Brecht, T.: Exploring the benefits of teams in multiagent learning. In: IJCAI (2022)
15. Radke, D., Larson, K., Brecht, T.: The importance of credo in multiagent learning. ALA Workshop at AAMAS (2022)
16. Radke, D., Orchard, A.: Presenting multiagent challenges in team sports analytics. In: AAMAS (2023)
17. Ritchie, R., Harell, A., Shreeves, P.: Pass evaluation in women's olympic ice hockey. In: Proceedings of the 5th International ACM Workshop on Multimedia Content Analysis in Sports. pp. 65–73 (2022)
18. SMT: SportsMEDIA technology. `https://www.smt.com` (2021), accessed: 2023-03-19
19. Spearman, W.: Beyond expected goals. In: Proceedings of the 12th MIT sloan sports analytics conference. pp. 1–17 (2018)
20. Spearman, W., Basye, A., Dick, G., Hotovy, R., Pop, P.: Physics-based modeling of pass probabilities in soccer. In: Proceeding of the 11th MIT Sloan Sports Analytics Conference (2017)
21. Steiner, S., Rauh, S., Rumo, M., Sonderegger, K., Seiler, R.: Outplaying opponents—a differential perspective on passes using position data. German Journal of Exercise and Sport Research **49**, 140–149 (2019)
22. West, R.M.: Best practice in statistics: Use the welch t-test when testing the difference between two groups. Annals of Clinical Biochemistry **58**(4), 267–269 (2021)
23. Zaccaro, S.J., Dubrow, S., Torres, E.M., Campbell, L.N.: Multiteam systems: An integrated review and comparison of different forms. Annual Review of Organizational Psychology and Organizational Behavior **7**(1), 479–503 (2020)

Simple and Practical Goal Importance Metrics for Ice Hockey

Rasmus Säfvenberg, Niklas Carlsson, and Patrick Lambrix

Linköping University, Sweden
firstname.lastname@liu.se

Abstract. To capture that not all goals are of the same importance, a new performance metric called the Game Points Importance Value (GPIV) was recently proposed. While this metric takes into account the expected impact that a goal has on the outcome of a game based on the context when the goal was scored, it relies on a relatively fine-grained state space. To address this problem, this paper presents simplified and more practical variations of the GPIV metric. Motivated by our analysis of the relative importance of different dimensions of the state space, we present two metrics that capture the most important component(s) of GPIV. Our evaluation shows that the metrics are relatively stable and capture most of the relative differences between GPIV and traditional metrics (e.g., goals, assist, points, and +/-). These results suggest that these simple and practical metrics are intuitive, capture most of the desirable variations that GPIV captures, and that the value of a goal can be well estimated using GPIV data based on historic data.

1 Introduction

In ice hockey, not all goals are of equal importance or have the same impact on the outcome of a game. For example, a game-tying goal in the final minute of regulation has a greater impact on the game outcome than a goal scored while leading by seven goals, as in the latter case the outcome is all but decided. In recent work [7, 8], we proposed a metric to quantify the importance of a goal on the game outcome in the National Hockey League (NHL). This metric, referred to as Game Points Importance Value (GPIV), accounts for the current goal difference (GD), manpower difference (MD), and the time of the game when the goal was scored. For each such (goal) state, we then compute a GPIV value, quantifying the goal importance, as the estimated change in the weighted probabilities (before vs. after the goal) for winning, losing in overtime, and losing in regulation. As an example, a game-tying goal in the final minute of regulation will increase the probability of winning and losing in overtime while simultaneously reducing the probability of losing. Such goal will therefore obtain a relatively large GPIV. In contrast, a goal scored when leading by a large margin will have negligible impact on the expected outcome, resulting in a small GPIV.

One downside of GPIV and other complex metrics is that they rely on a relatively fine-grained state space. To address this problem, this paper presents and evaluates two simplified and more easy-to-use variations of the GPIV metric. To derive and motivate these metrics, we make use of a decision tree and estimates of the variations that each of the state parameters are responsible for when using the original GPIV metric, and then define approximate metrics based on the insights provided. The resulting approximations capture the most important component(s) of GPIV and provide practitioners with an easy and straightforward application of the metric to real-time situations.

Like the pure GPIV metric, instead of attributing each goal an equal value, the approximations assign each goal a value based on the current state. However, by using a much smaller set of states, the approximations provide a more intuitive description of which goals have the highest importance within a game. These simplified valuations of player performance can therefore provide fans, teams, and media with an easy-to-apply metric for evaluating and comparing players that account for goal importance. Our evaluation of the approximations also shows that the metrics are relatively stable, allowing past seasons (or past games played) to be used to estimate and apply the metrics on current and future games. As desired, the metrics also have stronger correlations with GPIV than with the corresponding traditional metrics, and the relative player ranking variations (compared to traditional metrics) capture most of the explainable variations that have been observed using GPIV. These results suggest that these simple metrics are practical, intuitive, and capture most of the desirable goal importance variations captured by GPIV.

Throughout the remainder of this paper, data from the 2013-2014 NHL regular season is used to illustrate the approach.

2 Two simplified GPIV metrics

The original GPIV metric is based on each goal being its own state, where a state is represented by a time (in seconds), GD, and MD. However, not all these components are of equal importance for the computation of the GPIV value.

To evaluate the importance of each variable (i.e., time, GD, and MD), a decision tree was fitted with GPIV as the outcome and time, GD, and MD as variables. These results are shown in Figure 1. From this decision tree, we observe that GD is the variable with the most splits, followed by time, while MD had no splits. This relative ranking is also echoed by the variable importance (summarized in the same figure): GD (216.177) being the most important variable, time (76.586) being the second most important variable, and MD (0.398) being the least important variable. The variable importance is computed by summing the contribution of each variable (either as a primary or surrogate splitter) with a higher value corresponding to a higher contribution.

Looking closer at the decision tree, we also note that the least important goals were scored while already in the lead ($GD \geq 1$) within the first two periods, while the most important goals were goals in the final five minutes of regulation while trailing by one goal.

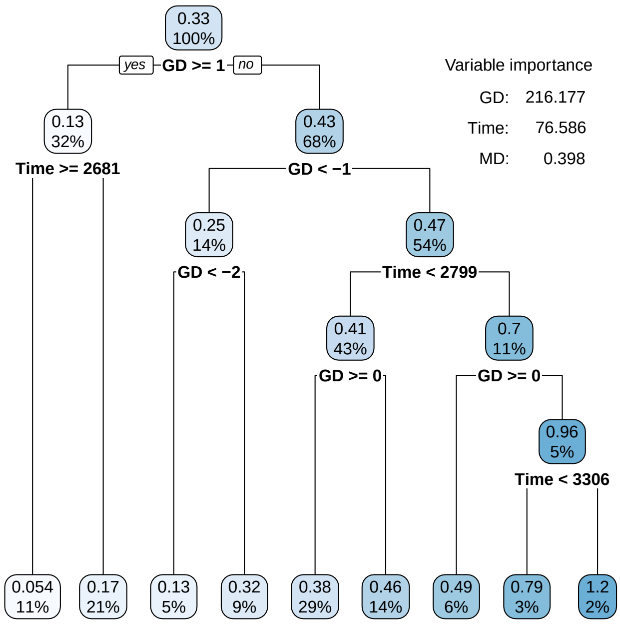


Fig. 1: Decision tree with GPIV as the outcome and GD, MD, and time as variables. Darker colors represent higher GPIV values, with the percentage describing the number of goals in each node.

These findings suggest that simplified metrics can be created by grouping some subsets of GD cases and time cases into a smaller set of categories, and that MD can be ignored as its importance is far smaller than the other factors.

First, for the GD dimension, we identified the following primary classes: reducing the deficit (GD ≤ -2), tying the game (GD $= -1$), taking the lead (GD $= 0$), and extending the lead (GD ≥ 1). For completeness, we also include the special case of an overtime (OT) winning goal, which exclusively occurs in overtime. Table 1 summarizes the average GPIV scores for each of these goal types (i.e., when considering GD only). Here, we note that the relatively big differences in the average GPIV of tying the game (0.594) and extending the lead (0.125), highlighting the value of such differentiating metric.

Second, for the time dimension, we selected to group goals according to the period a goal was scored but always considered them in combination with the above GD categories. The average GPIV values for each of these combined categories are shown in Table 2.

Based on the above categorizations, we can then define the two simplified GPIV metrics: *GD only* and *GD+Period*. For our notation, we let $\mathrm{GPIV}^*_{\mathrm{GD\ only}}$ be the approximated GPIV using GD only, and $\mathrm{GPIV}^*_{\mathrm{GD+Period}}$ be the approximated GPIV using both GD and period. In both cases, we simply assign each goal in a category the average GPIV value of all goals of that type. In the case

Table 1: Average GPIV per goal type when considering GD only.

Situation	Average GPIV	Goals
Reducing the deficit	0.249	887
Tying the game	0.594	1,263
Taking the lead	0.396	2,184
Extending the lead	0.125	2,102
OT winner	0.500	129

Table 2: Average GPIV per goal type when considering both GD and period. An OT winner has an average GPIV of 0.5.

	First period		**Second period**		**Third period**	
Situation	Average GPIV	Goals	Average GPIV	Goals	Average GPIV	Goals
Reducing the deficit	0.241	83	0.294	368	0.212	436
Tying the game	0.424	317	0.451	493	0.868	453
Taking the lead	0.370	1099	0.392	646	0.465	439
Extending the lead	0.180	397	0.162	789	0.070	916

that such GPIV values are based on prior games or seasons, these approximate GPIV values can therefore quickly be calculated at the time that a goal is scored.

When only considering GD, we note that tying the game is the most important goal, followed by OT winner, taking the lead, and reducing the deficit. Extending the lead is the least valuable goal. Similar conclusions can be drawn when also considering the period of the goal, although goals that tie the game or take the lead have an increased value as the game progresses. On the contrary, goals that reduce the deficit are most important in the second period while goals that extend the lead are most important in the first period, with the least important goals occurring in the third period for both situations. Another observation is that both simplifications lead to all goals having a positive value, which need not be the case in the full GPIV implementation [7, 8].

3 Stability of metrics

For previous estimations of the weights given to each goal to be useful, the metric should not change too drastically. Figures 2 and 3 visualize how the GPIV weights vary over time (on a season-per-season basis) for the approximate GPIV metrics based on GD only and GPIV based on GD+Period, respectively. When only considering GD, the GPIV weights exhibit low variability with stable weights over time. We also observe a strict order of the relative importance of the type of goals (matching the importance order from Table 1): goals that tie the game are the most important, followed by goals taking the lead, goals reducing the deficit, and goals extending the lead.

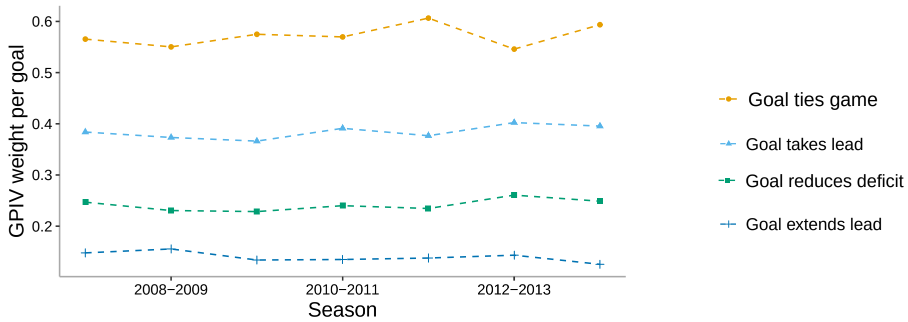


Fig. 2: GPIV weights by season when considering GD only.

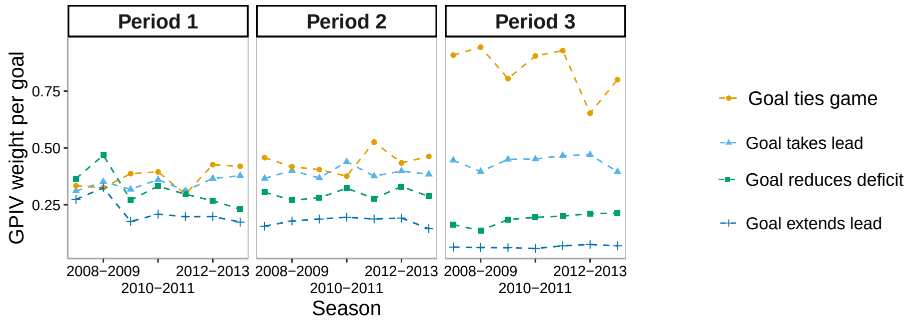


Fig. 3: GPIV weights by season when considering GD+Period.

If we also account for the period in which the goal occurs additional insights can be found. In general, regardless of the period, the least important goals are the goals that extend the lead, where the importance decreases with time. Although goals that reduce the deficit are the third most important goal in both the second and third periods as well as most seasons, they had the highest importance in the first period in both the 2007-2008 and 2008-2009 seasons. The importance of these goals also decreases in the third period. Another overall trend present is that goals taking the lead are the second most important goal type, with a mainly stable weight across all periods. The most important goals are found when tying the game in the third period, with the importance of a game-tying goal increasing as the game progresses.

4 Evaluation

Ideally, we would like the approximate metrics to capture most of the performance variations observed with GPIV. To determine if the metrics behave in a similar way as GPIV, we performed correlations comparisons and a rank-based analysis in which we compare with both GPIV and the corresponding traditional metrics. Some of these results are presented next.

Table 3: Spearman correlation between GPIV, simplified GPIV, and traditional metrics. Blank cells indicate a correlation between the same metric.

			Simplified GPIV metrics	
Class	Prior metric	Full GPIV	GD only	GD+Period
Goals	Traditional	0.971	0.979	0.975
Goals	Full GPIV		0.989	0.993
Assists	Traditional	0.979	0.986	0.983
Assists	Full GPIV		0.993	0.996
Points	Traditional	0.988	0.992	0.990
Points	Full GPIV		0.996	0.998
+/-	Traditional	0.775	0.843	0.802
+/-	Full GPIV		0.924	0.962

4.1 Correlation comparisons

As a point of comparison, we compute the Spearman rank correlation between the GPIV-based and the traditional metrics of goals, assists, points, and +/-. The correlations can be found in Table 3. Here, the class indicate what type of action the metric is calculated (i.e., the goals scored by a player, the assists made by the player, the sum of the first two, and whether a player was on the ice or not when a goal was scored during even strength). For each class we then present the correlation between the traditional metrics of that class and the three corresponding GPIV metrics (first row of each class), as well as between the full GPIV metric and the two approximations (second row of each class). As an example, we compute the correlations between Traditional Goals and Full GPIV by considering the total seasonal values for all players (Goals and GPIV-G).

First and most importantly, we note that the correlations between the full GPIV and the simplified GPIV metrics (second row for each class) are higher than the correlation with the traditional metrics (first rows). This suggests that the simplified metrics capture the most important variations of the full GPIV.

The table also highlights that GPIV has the lowest correlation of the considered methods, with both approximate methods having a higher correlation for all metrics. In particular, the simplified GPIV based on GD only was observed to have the highest correlation of all methods. This can be explained by GPIV considering a larger number of possible states, where some goals receive little to no value, which in turn lowers the correlation as the contrast between a goal of value one and close to zero is far larger than the approximate methods where the lowest value is 0.125 (simplified GPIV based on GD only) and 0.07 (simplified GPIV based on GD+Period).

Similarly, Figure 4 depicts the correlation for each pair of metrics across the analyzed seasons. We observe that goals, assists, and points all exhibit similar patterns over time, while +/- differs from the rest, particularly in the 2012-2013

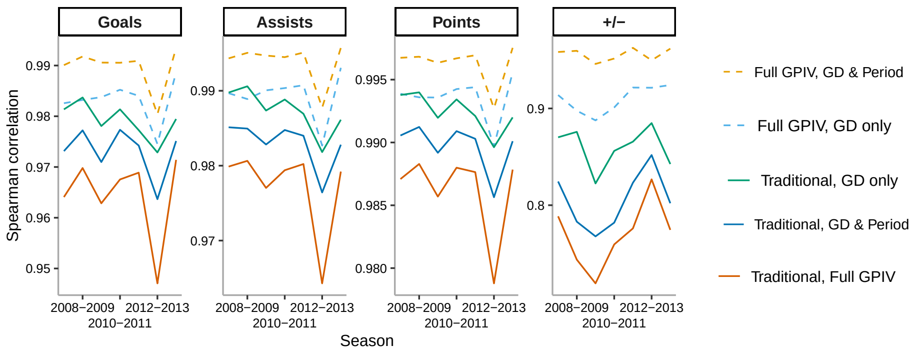


Fig. 4: Spearman correlation for each pair of metrics by season.

Table 4: Top-10 players for GPIV-P for the 2013-2014 season. Change is the difference in rankings for traditional and GPIV.

Rank							
P	GPIV-P	Change	Player	Position	P	GPIV-P	GPIV-P/P
1	1	0	Sidney Crosby	C	104	36.360	0.351
8-11	2	6	Alex Ovechkin	R	79	30.415	0.385
8-11	3	5	Nicklas Bäckström	C	79	29.199	0.370
19-22	4	15	Blake Wheeler	R	69	29.114	0.422
8-11	5	3	Joe Pavelski	C	79	27.995	0.354
4	6	-2	Tyler Seguin	C	84	27.614	0.329
3	7	-4	Claude Giroux	C	86	27.440	0.319
19-22	8	11	Kyle Okposo	R	69	26.951	0.391
16-18	9	7	Anze Kopitar	C	70	26.327	0.376
6-7	10	-4	Phil Kessel	R	80	26.225	0.328

season, which can be attributed to the lockout. Moreover, the correlations for +/- have a larger range, between 0.72 and 0.96, while goals, assists, and points all have values between 0.94 and 0.99. As desired, the strongest correlation was consistently observed between the full GPIV metric and its approximations. Among the five pairs, the correlation between the traditional and full GPIV metrics was the lowest, although the correlation is still high.

4.2 Player rankings

GPIV can also be used in the context of player valuation. In this section, we provide the top-ten rankings for GPIV points (P) for each of the three methods for comparison. The top-ten players according to the GPIV method can be found in Table 4, while Tables 5 and 6 contain the results for the simplified GPIV based on GD only and (simplified GPIV based on GD+Period), respectively. Overall, eight players are present in all three tables: Sidney Crosby, Alex Ovechkin, Nick-

Table 5: Top-10 players for simplified GPIV*-P based on GD only for the 2013-2014 season. Change is the difference in rankings for traditional and GPIV*.

Rank								
P	GPIV*-P	GPIV-P	Change	Player	Position	P	GPIV*-P	GPIV*-P/P
1	1	1	0	Sidney Crosby	C	104	36.520	0.351
8-11	2	5	6	Joe Pavelski	C	79	28.546	0.361
19-22	3	4	16	Blake Wheeler	R	69	28.514	0.413
8-11	4	2	4	Alex Ovechkin	R	79	27.742	0.351
2	5	11	-3	Ryan Getzlaf	C	87	27.406	0.315
4	6	6	-2	Tyler Seguin	C	84	27.300	0.325
8-11	7	3	1	Nicklas Bäckström	C	79	27.254	0.345
6-7	8	10	-2	Phil Kessel	R	80	27.185	0.340
3	9	7	-6	Claude Giroux	C	86	26.643	0.310
6-7	10	13	-4	Taylor Hall	L	80	26.399	0.330

Table 6: Top-10 players for simplified GPIV*-P based on GD+Period for the 2013-2014 season. Change is the difference in rankings for traditional and GPIV*.

Rank								
P	GPIV*-P	GPIV-P	Change	Player	Position	P	GPIV*-P	GPIV*-P/P
1	1	1	0	Sidney Crosby	C	104	36.485	0.351
8-11	2	2	6	Alex Ovechkin	R	79	28.877	0.366
8-11	3	5	5	Joe Pavelski	C	79	28.129	0.356
19-22	4	4	15	Blake Wheeler	R	69	28.072	0.407
8-11	5	3	3	Nicklas Bäckström	C	79	27.985	0.354
6-7	6	10	0	Phil Kessel	R	80	26.787	0.335
4	7	6	-3	Tyler Seguin	C	84	26.751	0.318
13	8	12	5	Joe Thornton	C	76	26.749	0.352
3	9	7	-6	Claude Giroux	C	86	26.687	0.310
2	10	11	-8	Ryan Getzlaf	C	87	26.264	0.302

las Bäckström, Blake Wheeler, Joe Pavelski, Tyler Seguin, Claude Giroux, and Phil Kessel. Although the rankings of players differ between the tables, Sidney Crosby remains atop all three tables with similar GPIV values. A possible explanation is the large difference in points between him and the lower-ranked players. If we consider the simplified GPIV based on GD only, simplified GPIV based on GD+Period, and GPIV in order of complexity, we can also see that both Alex Ovechkin and Nicklas Bäckström gain ranks with increased GPIV complexity. Their ascent is likely a result of the number of goals scored in close games, losing by one or tied, at the end of the game or in overtime with Ovechkin as the likely goalscorer and Bäckström with the assist. Similarly, Blake Wheeler also gains ranks in all cases while also having the highest average importance per point. On the contrary, Ryan Getzlaf, who ranked second in total points, loses ranks as he is ranked fifth for the simplified GPIV based on GD only, tenth for the GD

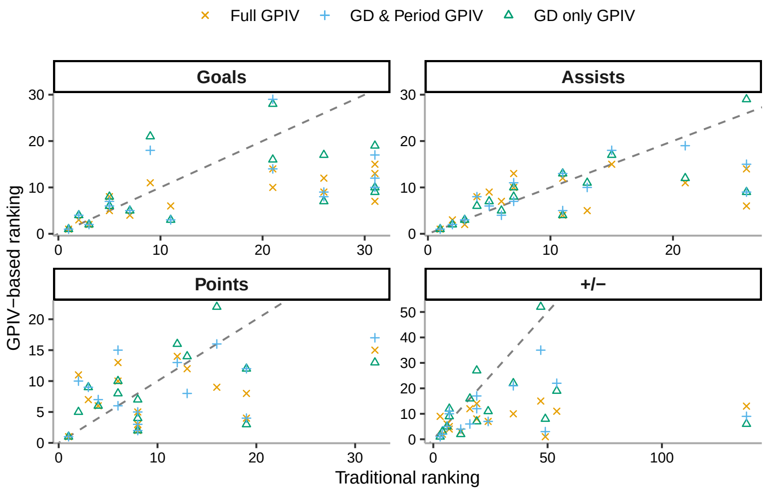


Fig. 5: Top-15 ranked players per full GPIV metric. Note that some players share the same rank on the x-axis.

and Period simplification, and eleventh for the full GPIV. These observations are consistent with the correlation results.

To investigate the impact that each GPIV-based method has on play ranking, Figure 5 shows ranking using each GPIV-based metric (y-axis) vs. the traditional ranking (x-axis) for the top-15 ranked players using the full GPIV metric. For both GPIV-Goals and GPIV-Assists it appears that the top-ranked players are mainly unaffected by the different methods as their ranks remain stable. The variability in ranking becomes more prevalent for somewhat lower traditional rankings, as the highest climbing players may have higher variability in ranking across the GPIV-based methods. As an example, Anze Kopitar was ranked 28th and 29th when considering the GPIV-based approximations based on GD only and GD+Period, respectively, but for the full GPIV he was ranked 14th. For GPIV-Points some players also manage to maintain a stable rank, for instance, the top-ranked player Sidney Crosby. The GPIV-based ranking for points also had the lowest range, between 1 and 22. On the contrary, +/- had the highest range, between 1 and 52, while also having the largest spread for traditional ranking with Matt Duchene climbing from 137th in the traditional +/- to top 15 in GPIV-+/-. The variability of rankings for the GPIV-based +/- metrics is also the highest, as some players have a larger discrepancy between their rankings of the different methods. For instance, Sidney Crosby had rankings of 15th (full GPIV), 35th (GPIV based on GD+Period), and 52nd (GD only).

Another way to illustrate the difference in ranking between the traditional ranking and the GPIV ranking can be found in Figure 6, where the top-15 players, for each full GPIV metric, are visualized. Here we note that the largest rank increases are found for +/-, with Matt Duchene gaining over 100 ranks

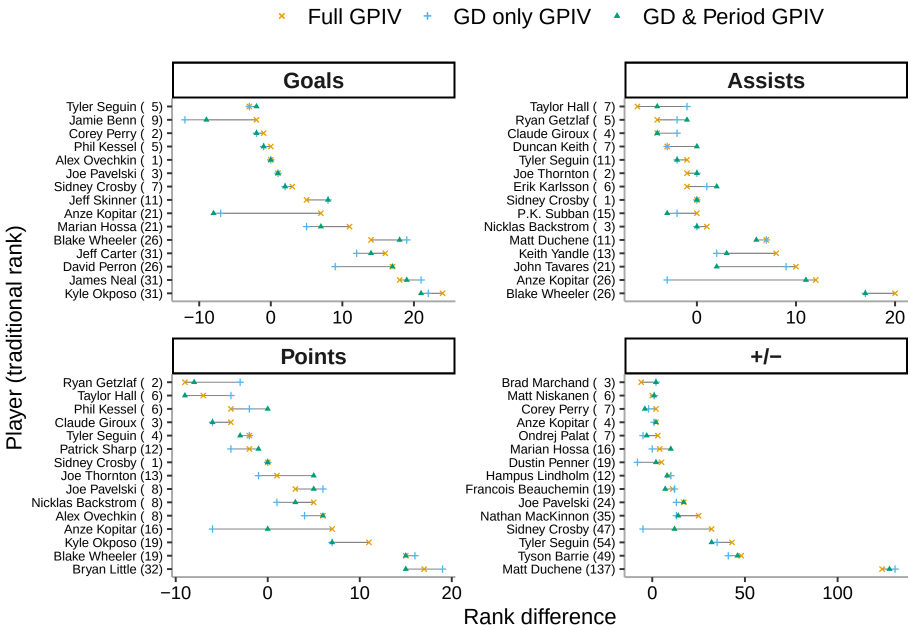


Fig. 6: Rank difference between the traditional and GPIV-based ranking for the top-15 per full GPIV metric.

while e.g., Tyler Seguin and Tyson Barrie also gain approximately 50 ranks. Another thing to note is that for GPIV-+/-, few players lose ranks compared to the traditional +/-. As for goals, assists, and points, players losing ranks is more prevalent. For GPIV-Goals we also observe that the different GPIV-based methods result in mostly similar rankings, although Jamie Benn and Anze Kopitar have a larger difference between the full GPIV and the approximations. In addition, for the simplified GPIV-based methods, Anze Kopitar loses ranks for GPIV-Goals while he gains ranks for the full GPIV. The same results can also be seen for Anze Kopitar in GPIV-Assists and GPIV-Points. In general, GPIV-Assists and GPIV-Points also have a higher difference in rankings when comparing the GPIV methods than goals. We also note some of the big climbers, e.g. Bryan Little (Points) and Blake Wheeler (Assists and Points), while players who lose ranks, e.g., Ryan Getzlaf (Assists and Points) and Taylor Hall (Assists and Points), also have variability between the different GPIV-based methods, with the full GPIV tending to assign them the lowest ranking.

5 Related work

The most used performance metrics in ice hockey are the total number of goals, assist, and points accumulated over a season or some other time period or set of games. Like these metrics, the GPIV metric and the GPIV-based approximations presented here are calculated as the sum of all goals. The main difference is in the weight given to each goal and that these traditional metrics do not account for the potential impact a goal may have on the game outcome.

Some extensions to these traditional metrics have been proposed (e.g., for the +/- metric [11, 4]) and combined metrics have been proposed (e.g., based on principal component analysis [5]). Others have proposed player performance metrics take game context into account (e.g., the probability that an event leads to a goal in the subsequent 20 seconds [15]) or that incorporate the game models using Markov games where two opposing sides (i.e., the home team and the away team) try to reach states in which they are rewarded (e.g., scoring a goal) [18, 6, 13, 16, 17, 9, 14, 10]. One critique of these more advanced metrics is that they are not easily understandable by or explainable to practitioners such as coaches, players, and GMs. In this work, we aim to present such simpler and more practical metrics that still differentiate between the potential impact that a goal may have on the game outcome.

Prior works have also considered the importance of scoring the first goal [1], a two-goal lead [2], and late-game reversals [3]. For example, it was found that teams that take a two-goal lead win in 83% of games, while having the lead after two periods leads to a win in 84% and 80% of games for the home and away team, respectively.

Except our prior work defining the original GPIV metric, the only other work that considers the importance of goals is the added goal value (AGV) metric presented by Pettigrew [12]. The importance of a goal is based on GD and time and is defined using win probabilities for that context and neighboring contexts (with GD one higher and one lower). The AGV is then defined for a player by comparing the importance of the player's goals to the importance of all other players' goals.

Some players can have a positive (or negative) impact even when they are not the player scoring the goal or assisting to the goal. Perhaps the most used metric to estimate the value a player brings to team performance (during 5-on-5 play) is the +/- metric. While the metric has been criticized due to its disregard of contextual information [18], alternative approaches typically also ignore the importance of individual goals. Interesting examples falling in this category include works based on hazards models [18], regularized logistic regression for predicting player impact on scoring [4], or models that sum over all actions performed by a player [13] or set of players [10] when on the ice at the same time.

6 Conclusions

This paper has presented two approximate GPIV metrics: GD only and GD+Period. The design of the metrics was motivated by our analysis of the relative importance of different dimensions of the state space, and our evaluation demonstrated that the metrics are relatively stable and capture most of the relative differences between GPIV and traditional metrics (e.g., goals, assist, points, and +/-). The presented metrics are practical, intuitive, capture most of the desirable variations that GPIV captures, and show that the value of a goal can be well-estimated using GPIV data based on historic data. These properties should make it desirable for fans, teams, and media that want an easy-to-apply metric for evaluating and comparing players that account for goal importance.

References

1. Brimberg, J., Hurley, W.: A note on the importance of the first goal in a National Hockey League game. International Journal of Operational Research **6**(2), 282–287 (2009)
2. Brimberg, J., Hurley, W.: Do professional hockey teams with a two-goal lead lose more often than they should? International Journal of Operational Research **15**(2), 226–233 (2012)
3. Gill, P.S.: Late-game reversals in professional basketball, football, and hockey. The American Statistician **54**(2), 94–99 (2000)
4. 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)
5. 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)
6. 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)
7. Lambrix, P., Carlsson, N.: Performance metrics for ice hockey accounting for goal importance. In: Linköping Hockey Analytics Conference. pp. 11–25 (2022)
8. Lambrix, P., Carlsson, N., Säfvenberg, R.: Goal-based performance metrics for ice hockey accounting for goal importance (2023), submitted
9. 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)
10. Ljung, D., Carlsson, N., Lambrix, P.: Player pairs valuation in ice hockey. In: Machine Learning and Data Mining for Sports Analytics. MLSA 2018. pp. 82–92 (2019)
11. Macdonald, B.: A Regression-Based Adjusted Plus-Minus Statistic for NHL Players. Journal of Quantitative Analysis in Sports **7**(3) (2011)
12. Pettigrew, S.: Assessing the offensive productivity of NHL players using in-game win probabilities. In: MIT Sloan Sports Analytics Conference (2015)
13. Routley, K., Schulte, O.: A Markov Game Model for Valuing Player Actions in Ice Hockey. In: Uncertainty in Artificial Intelligence. pp. 782–791 (2015)
14. Sans Fuentes, C., Carlsson, N., Lambrix, P.: Player impact measures for scoring in ice hockey. In: MathSport International 2019 Conference. pp. 307–317 (2019)
15. 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)
16. 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)

17. 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)
18. 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)

The Importance of Special Teams in Ice Hockey

Rasmus Säfvenberg[1], Mikael Svarén[2], Niklas Carlsson[1], and Patrick Lambrix[1]

[1] Linköping University, Sweden
[2] Dalarna University, Sweden

Abstract. This paper explores the significance of special teams, particularly powerplay, in ice hockey. Despite the commonly held perception of their importance, little research has examined the impact of powerplay and penalty kill performance on overall team success. The paper uses several seasons of NHL data to characterize goal-scoring and manpower opportunities, and perform analysis from several perspectives. The results indicate that individual even strength goals and powerplay goals have similar value, but the larger share of even strength goals scored over a season makes even strength play a more important contributor to team success. The paper also finds a high correlation between teams that perform above/below average during even strength and powerplay. This study provides insights into the dynamics of ice hockey gameplay and the role of special teams in determining team success.

1 Introduction

Like most games, ice hockey is played according to a set of rules, and if a player violates a rule during the game, the team responsible for the violation is given a penalty. Furthermore, the player who committed the violation is then sent to the penalty box, and as a result, the opposing team is typically given a temporary manpower advantage to play against the penalized team.

A penalty in ice hockey can significantly alter the game's dynamic by disrupting the offensive and defensive strategies of both teams. The non-penalized team gains a numerical manpower advantage, which typically leads to increased possession of the puck closer to the opposing team's net, providing them with a boost in offensive capabilities [9]. On the other hand, the penalized team is often forced to play more defensively to prevent the non-penalized team from scoring while the player is in the penalty box.

Whenever a penalty occurs in ice hockey, both teams field their "special teams"; i.e. their powerplay unit or their penalty killers. The team who receives a numerical advantage from the penalty goes on the powerplay and typically play their strongest offensive players in an attempt to maximize their scoring chances, while the penalized team play their best defensive players in an attempt to prevent the other team from scoring during the penalty [1].

Due to the increased scoring opportunities that come with powerplay opportunities, both teams and fans often put great weight on the powerplay. However, the perceived importance must also be put in perspective of the full game and the impact that powerplay goals have on the outcome of the games. Here, it should be noted that NHL games (studied here) span 60 minutes and there is no guarantee that a given team will be on the powerplay. Instead, the vast majority of the game is played in even strength.

Although commonly perceived to be a vital part of team success, the importance of special teams, in particular powerplay, has not been extensively studied. Research by [10] has reported that gaining a powerplay opportunity can drastically increase the conditional probability of scoring a goal. Similarly, a higher goal scoring probability while having a numerical advantage was described in [6]. Another area that has not been thoroughly explored is the relationship between team success in special teams and overall team success, in particular, the exact dynamics of how performance in powerplay and penalty kill influence even strength performance and vice versa.

This paper studies and quantifies the importance of special teams in ice hockey from several perspectives, considers its contribution to team success, and compares the importance to the importance of even strength team success. After describing our dataset (Section 2) and characterizing the goal scoring and manpower opportunities in a typical game (Section 3), we perform our analysis from several perspectives. In particular, we use the recent GPIV metric to analyze the importance of individuals goals (Section 4), use correlations (Section 5) and model-based evaluations (Section 6) to consider how much powerplay goals contribute to the outcome of individual games, and finally we study correlations between team success on a per season basis and perform above/below average at even strength and on the powerplay (Section 7).

Our results show that individual even strength goals and powerplay goals have similar value (when scored at similar times of the game and when the goal differential is the same), but that the larger share of even strength goals (compared to powerplay goals) scored over a season makes even strength play a more important contributor to team success. However, it should also be noted that we have found high correlation between teams that are above/below average during even strength and during powerplay, suggesting that in many cases the teams that have above/below average players for even strength also have above/below average players for powerplay.

2 Dataset

This paper uses data from `https://www.NHL.com` and their public API. We use the information regarding penalty and goal statistics. More specifically, the duration of various manpower situations and whether one or more goals were scored during these situations. The seasons included in the data are the regular seasons from 2010-2011 to 2021-2022 where overtime periods were excluded. Here, we note that the 2012-2013 season consisted of only 48 games due to a lock-out, while the 2019-2020 and 2020-2021 seasons were shortened due to COVID-19.

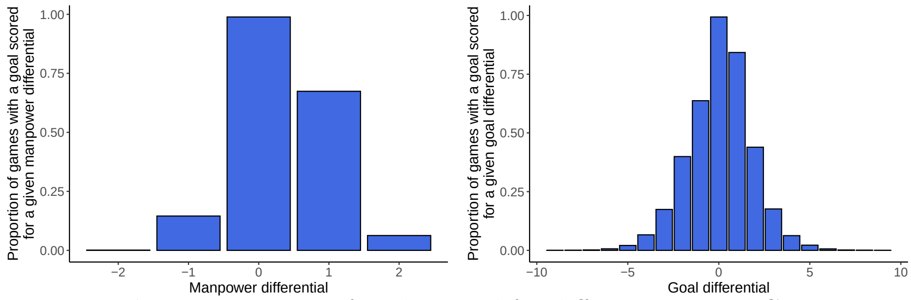


Fig. 1: Frequency of goals scored for different MD and GD.

3 Goal scoring and manpower opportunities

3.1 Goal scoring

While almost all games include even strength goals, less than 75% of games include at least one powerplay goal. This is illustrated in Figure 1, where we show the proportion of games with at least one goal scored with different manpower differentials (MD) from the scoring team's point of view. For completeness, we also include the same statistics for the goal differential (GD) i.e., the goals scored by the scoring team minus the goals scored by the other team. We note that in 99% of cases, a goal is scored for GD = 0 in games with a goal scored in regulation and/or overtime, while a similar value can also be observed for MD = 0. For GD, there is also a slightly higher prevalence for scoring when leading by one goal, compared to trailing by one. Empty net goals may be a factor here. Otherwise, there is a balance between goal scoring in the case of a positive and negative GD of the same absolute quantity. When considering MD, we note that many games have goals scored while having a numerical advantage, particularly when playing 5v4 or 4v3. It is also noteworthy that scoring goals in 5v3 (i.e., MD = 2) is less common than scoring while shorthanded (e.g., MD = -1).

3.2 Frequency of manpower scenarios

Figure 2 shows the proportion of each manpower scenario, and how it has changed over time. We note that most of the game is played in 5v5 (approx. 75% to 80% of the total game time), while either team having a one-man advantage occurs between 15% to 19% of the game. Moreover, around 1-2% of the total time is played 4v4, with the remaining time distributed for a two-man advantage and 3v3. We note a slight increase in the fraction of time spent in 5v5 from 2010-2011 to 2021-2022.

3.3 Powerplay scoring in different manpower scenarios

Naturally, the more opportunities a team obtains on the powerplay, the greater chance there is that a team scores at least one powerplay goal. Here, we quantify

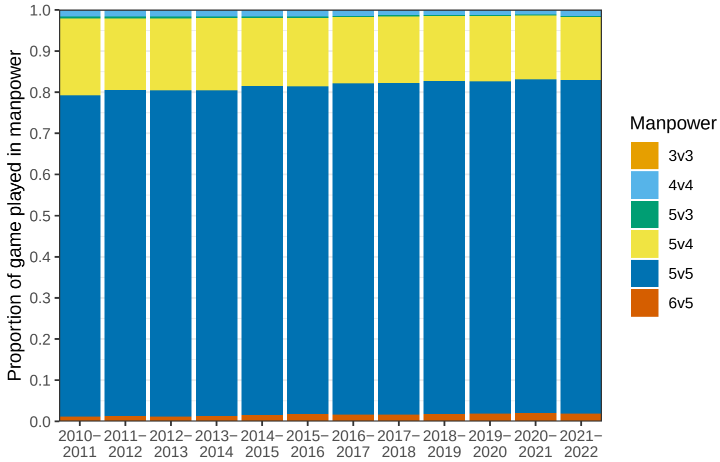


Fig. 2: Proportion of each manpower situation over time.

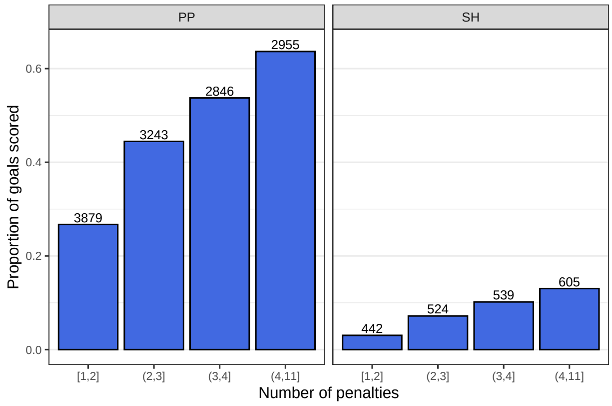


Fig. 3: Proportion of goals scored by number of penalties. The number on the bar shows the total number of goals.

the fraction of a game when a team scored at least one powerplay (or shorthanded) goal as a function of the number of powerplay (or shorthanded) opportunities they obtain in a game. These results are shown in Figure 3. For completeness, we include the number of games associated with each case. We note that the fraction of games that at least one powerplay goal was scored goes up from 25% when only having one or two powerplay opportunities during a game to approximately 44% when having three penalty opportunities, 54% when four penalties, and 64% when five or more penalties.

3.4 Goal scoring during double-minors and major penalties

Although the two-minute minor penalty is the most common type of penalty, double-minors and majors (including match penalties) may still occur and affect the state of the game. A double-minor (2+2 minutes) lasts four minutes at most,

Table 1: Long penalties scoring rate.

	Double-minor		Major	
Goals scored	Occurrences	Proportion	Occurrences	Proportion
0	790	0.688	383	0.670
1	307	0.267	142	0.248
2	51	0.044	37	0.065
3 or 4			9+1	0.018

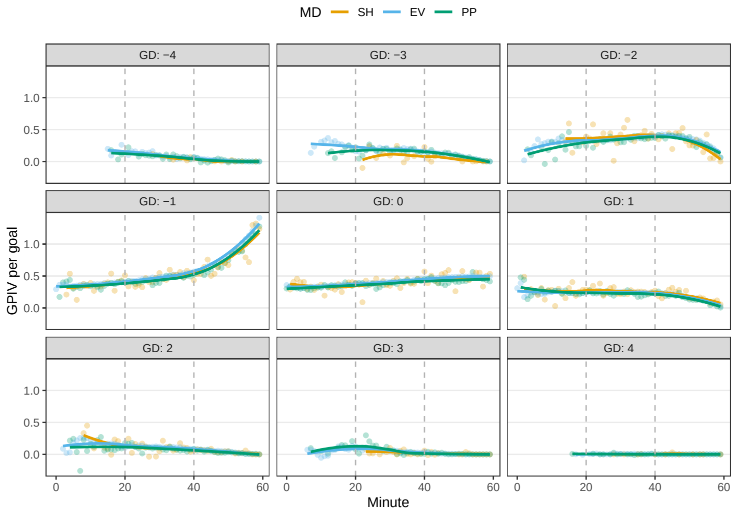


Fig. 4: GPIV per goal and minute.

although a goal will remove up to two minutes of penalty time and two goals will nullify it completely. In contrast, there is no upper limit for how many goals can be scored during a major penalty. A major penalty lasts five minutes and is not ended if a goal is scored. For the investigated seasons, there was a total of 1,148 double-minors and 572 majors that caused a manpower change. Coincidental penalties that cancel each other out are not included in this count. This way we exclude the case when two players from opposing teams draw a major penalty each for fighting. The observed outcomes and scoring rates for double-minors and majors can be found in Table 1. We note that 68.8% and 67% of double-minors and major penalties, respectively, end without a goal being scored. However, a few double-minors have two goals (4.4% of cases). Similarly, of the major penalties, only 8.2% result in two or more goals being scored.

4 The relative value of individual goals

For this analysis, we use the GPIV metric [4,5] to assign every goal an importance value that take into account the goal state, defined by the time, GD, and MD at the time that the goal was scored. Intuitively, the GPIV can be seen as a

Fig. 5: GPIV per goal. The numbers above bars indicate the total number of goals.

weighted metric that captures the change in the probabilities (before vs. after the goal) for winning, losing in overtime, and losing in regulation, as well as the number of points that a team obtains if one of those game outcomes take place.

In Figure 4 we illustrate how the average goal is valued for different GD and MD scenarios per minute. A general observation is that MD does not seem to affect goal importance, rather it is more reliant on GD and elapsed time. For instance, the highest average goal importance can be seen near the end of regulation when trailing by one goal, as these goals tie the game and result in a higher probability of attaining game points. Similarly, we note that a larger absolute GD (e.g., -4, -3, 3, and 4) tends to have negligible importance on the outcome of the game. However, the manpower curves suggest that there is no clear distinction between shorthanded, even strength, and powerplay.

Another aspect worth considering is the average GPIV per goal for different GD and MD and whether there are any noticeable differences between them. These comparisons are shown in Figure 5. Here it can be noted that the importance of a goal is the highest when trailing by one, and higher absolute GD are seldom important. When contrasting the different manpower scenarios we note that, in general, the goals scored in even strength have a higher average GPIV for a given GD despite a far larger sample size than shorthanded and powerplay goals. Interestingly, powerplay goals are typically the least important for a given GD, although the differences are somewhat small.

5 Impact on winning a game

To examine the importance of special teams scoring within a single game, we next consider the correlation of the scoring rates in each state with each game

Table 2: Spearman correlations.

Manpower	Game points	Win
EV For	0.524	0.507
EV Against	-0.559	-0.507
PP For	0.222	0.212
PP Against	-0.127	-0.112
SH For	0.115	0.112
SH Against	-0.235	-0.212

Table 3: Average goals for and against per game, by game points.

	Goals for				**Goals against**			
Game points	Total	EV	PP	SH	Total	EV	PP	SH
0	1.64	1.21	0.38	0.04	4.07	3.12	0.13	0.82
1	2.38	1.79	0.53	0.07	2.38	1.80	0.06	0.52
2	3.68	2.81	0.76	0.11	1.81	1.35	0.05	0.42

outcome. Naturally, the team with the most goals at the end of the game wins. However, does the amount of goals per manpower scenario have equal importance for this? To answer this, the Spearman correlation between winning a game (1 if win, 0 otherwise) and obtaining game points (2 if win, 1 if overtime loss, and 0 if loss) and the number of goals scored and allowed per game was investigated. The results are shown in Table 2.

We note that the strongest relationship between both winning and obtaining game points was found with even strength scoring, while powerplay goals for/shorthanded goals against ranked second, although with a weaker correlation. This result could be expected, as a majority of the game is played in even strength and we may therefore expect that most goals are scored during this manpower scenario.

These correlations can also be explained from the point of view of average scoring and conceding rates. These results are shown in Table 3. Here we note that, from both perspectives, most goals occur during even strength play while the average number of goals scored increases when the game points increase, while goals against decrease when game points increase. The same pattern can also be discerned for powerplay and shorthanded situations, as the scoring rates also increase with game points while goals against decrease with increased game points. Interestingly, if powerplay goals were to be excluded, most games would still have the same outcome.

6 Relationship between game points and manpower

With the inherent randomness in ice hockey, the goal importance in a given game may have larger variability than if considering the entire season. As an example, the powerplay efficiency in one game may be 100%, with the team scoring a goal in one powerplay opportunity, but this is not expected to be true for the entire season. Instead, the team success in powerplay typically ranges between 10% and 30%, where the best scoring teams average higher numbers than the worst scoring teams. Yet, only considering a team's powerplay efficiency without accounting for their skill in even strength and while shorthanded fails to fully contextualize the performance of a team. Therefore, to fully account for all of these situations, we implement several generalized additive models (GAMs)[3] models to investigate the relationship between game points and efficiency in various manpower situations. Table 4 summarizes these results. The choice of a GAM model is suitable as it allows for flexible modeling of the relationship between an outcome and a set of variables [11]. Here, each model has the same outcome, points accrued after a full season,[4] while the variables differ. To evaluate the out-of-sample quality of each model, data from 2010-2011 until 2020-2021 were used as training data while 2021-2022 was used as the test set. The variables in each model are:

- SH: Shorthanded goal differential per game.
- $\mathrm{SH_{te}}$: Interaction between shorthanded goals for and against per game.
- PP: Powerplay goal differential per game.
- $\mathrm{PP_{te}}$: Interaction between powerplay goals for and against per game.
- SP: Powerplay and shorthanded goal differential per game.
- $\mathrm{SP_{te}}$: Interactions between powerplay goals for and against per game, and shorthanded goals for and against per game.
- EV: Even strength goal differential per game.
- $\mathrm{EV_{te}}$: Interaction between even strength goals for and against per game.
- All: Interactions between even strength and powerplay goal differential per game, and even strength and shorthanded goal differential per game.
- $\mathrm{All_{te}}$: Interactions between even strength and powerplay goals for per game, and even strength and shorthanded goals allowed per game.

From the set of models, we note that the deviance explained (where a value of 0 indicates no explanation of the outcome while 1 provides a perfect explanation) varies by model, with specific manpower situations, i.e. shorthanded and powerplay, having the lowest value, while the model using all scenarios obtained the highest scores. These results indicate that merely considering a team's strength in, e.g., shorthanded or powerplay, is insufficient to explain the total team points they will accumulate over the season. By considering both shorthanded and powerplay situations in a model, we find increased deviance explained at 40-42%. In contrast, by only accounting for the quality of play in even strength situations we can explain approximately 85% of the deviance, and by including the other two scenarios this increases to 90%. Similarly, when evaluating the out-of-sample performance of the models, the even strength and all-inclusive model has

[3] With restricted maximum likelihood estimation and thin plate splines.

[4] For the non-82 game seasons, the accrued game points after their last game was generalized to an equivalent of 82 games.

Table 4: Model evaluation metrics.

Model	Res. Df	Res. Dev	Dev. Expl.	Training MSE	Test MSE
SH	330.9	58900.3	0.187	176.35	329.20
SH_{te}	328.5	58164.2	0.197	174.14	327.77
PP	331.5	55210.6	0.238	165.30	182.49
PP_{te}	330.0	54829.0	0.243	164.16	179.98
SP	331.0	43085.2	0.405	129.00	157.03
SP_{te}	324.2	41573.3	0.426	124.47	151.92
EV	331.7	11167.5	0.846	33.44	48.51
EV_{te}	327.5	10891.6	0.850	32.61	49.40
All	322.0	7212.0	0.900	21.59	30.17
All_{te}	324.7	7113.6	0.902	21.30	32.72

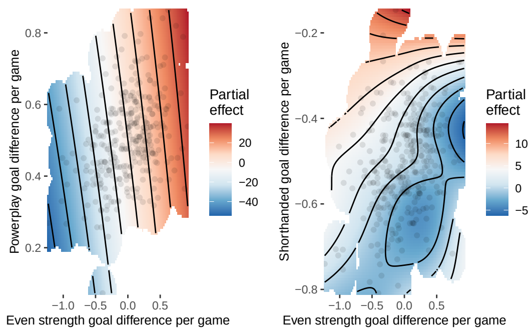


Fig. 6: Effect of manpower scenario goal differential on earning game points from the 'All' model. Partial effect of 0 means average, while red indicate more points and blue less points

the best performance, while the shorthanded and powerplay models have higher test MSE.

Thus, it becomes evident that the main component in explaining team success, measured in team points, lies in the quality of their even strength play. This stands in unison with the fact that a vast majority of an ice hockey game is being played in even strength, in particular during 5-on-5. A visualization of the best performing model, with respect to deviance explained and test MSE, can be seen in Figure 6. The figure shows the joint impact of two sets of variables: GD per game in EV and PP (left) and GD per game in EV and SH (right).

Overall, the model highlights results that are expected. In particular, a higher even strength goal differential typically leads to more game points, while higher goal differentials for powerplay and shorthanded also increase the expected number of game points.

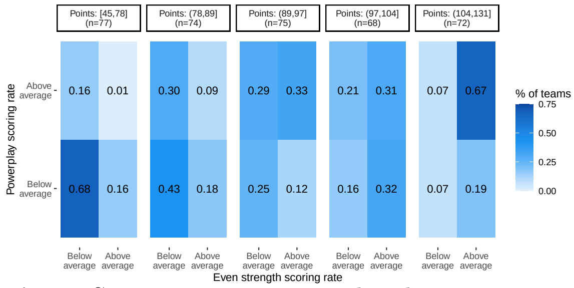


Fig. 7: Scoring rates compared to league average in PP and EV.

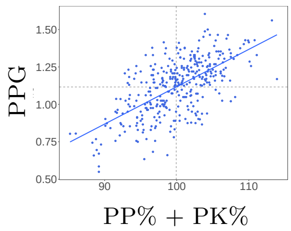


Fig. 8: Special teams (PP% + PK%) vs. points per game (all seasons).

7 Team-based per season analysis

7.1 Correlations

Even though even strength (as per the above results) explains most of the team success, it is clear that many good teams have a strong powerplay and many bad teams have weak powerplay. We expect that this is due to good teams having good players to put on the powerplay and bad teams often having to put weaker players on the powerplay. This is illustrated by the high concentration of the top teams (e.g., with more than 104 points in a season) having both above average even strength scoring rate and above average powerplay scoring rate (last cell in Figure 7) and the high concentration of the bottom teams (e.g., with less than 76 points in a season) having both below average even strength scoring rate and below average power-play scoring rate (first cell in Figure 7).

These observations can also be extended to special teams in general, where teams that can ice a strong 5v5 team often can ice a strong unit or two for both powerplay and shorthanded situations. For example, we have observed a strong correlation between the sum of the (PK% + SH%) and team success (Figure 8). We note that a team with (PK% + SH%) above 100% typically sees a net gain from special teams situations (assuming a similar number of powerplays and shorthanded situations) and teams with values below 100% generally are outperformed in special teams situations. When discussing penalty killing, it can also be noted that the goaltenders, who make up an important part of a strong penalty killing unit often also play a big part in a team's 5v5 success.

7.2 Longitudinal analysis

While there are exceptions (especially some teams becoming weaker), we have observed a relatively larger increase in the fraction of even strength (EV) goals compared to special teams goals (PP and SH). These statistics are shown on a team basis in Figure 9. The increase in goal-scoring in the figure aligns with an overall increase in scoring from 2010-2016 (2.71-2.79 goals per game) and 2017-2023 (2.94-3.18 goals per game).

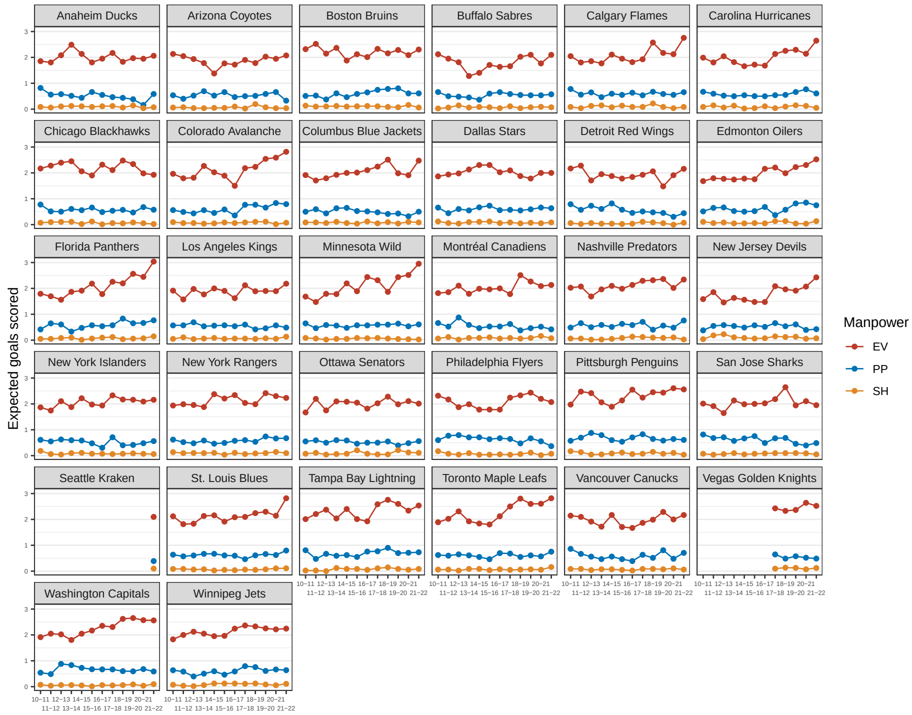


Fig. 9: Expected goals scored per team, season and manpower.

8 Discussion in a wider context

An important aspect is the impact the importance of special teams may have on player development among younger athletes [3]. What is viewed as important may be reflected in how the coach of a youth team chooses to coach, e.g. having a "win-first" mentality [12]. For instance, if powerplay is seen as a vital component of the game, it may affect teams at the youth level, where the focus is placed on specialized situations, e.g. powerplay and penalty killing, instead of focusing on fundamental individual skills, e.g. skating, passing, and stick-handling, and team skills, e.g. puck support and knowledge of tactical situations. A large proportion of practice may be dedicated to these specialized situations despite the majority of a hockey game being played in even strength [7]. Who gets to play in these specialized settings may also vary, as some coaches only select the best players while others allow most, if not all, players to participate. However, this typically changes when the stakes are higher, e.g. in playoffs or tournaments, where most coaches lean toward only choosing the players who they believe will maximize their winning chances [3]. This type of specialization may hamper individual development and affect both the preferred and non-preferred players negatively [8,2]. While our study cannot provide a clear answer to how much time kids should practice powerplay skills, it highlights that even strength play may have a greater impact on team success at the NHL level than powerplay.

9 Conclusion

In conclusion, our analysis of individual goal scoring, game outcomes, and team success across multiple seasons provides insight into the significance of powerplays and penalty kills in the NHL. We found that powerplay goals and even strength goals have similar values when scored at similar times of the game and when the goal differential is the same. However, even strength play is a more important contributor to team success due to the larger share of even strength goals scored over a season. Our results also show a high correlation between teams that perform well during even strength and powerplay, indicating that team success is closely linked to the overall skill level of the team's players.

Overall, our study highlights the importance of a team's ability to perform effectively on special teams, but also the importance of maintaining a strong even strength performance. The findings of this study may inform coaches and players on the relative importance of special teams versus even strength play and provide guidance for optimizing team strategies for success. Our study also suggests potential avenues for further research into the dynamics between special teams performance and overall team success in ice hockey.

References

1. von Allmen, P., Leeds, M., Malakorn, J.: Victims or Beneficiaries?: Wage Premia and National Origin in the National Hockey League. Journal of Sport Management **29**(6), 633–641 (2015)
2. Donnelly, P., Petherick, L.: Workers' playtime? child labour at the extremes of the sporting spectrum. Sport in Society **7**(3), 301–321 (2004)
3. Gilbert, W.D., Trudel, P.: Role of the coach: How model youth team sport coaches frame their roles. The Sport Psychologist **18**(1), 21–43 (2004)
4. Lambrix, P., Carlsson, N.: Performance metrics for ice hockey accounting for goal importance. In: Lambrix, P., Carlsson, N., Vernblom, M. (eds.) Linköping Hockey Analytics Conference. pp. 11–25 (2022)
5. Lambrix, P., Carlsson, N., Säfvenberg, R.: Goal-based performance metrics for ice hockey accounting for goal importance (2023), submitted
6. Pettigrew, S.: Assessing the offensive productivity of nhl players using in-game win probabilities. In: MIT Sloan Sports Analytics Conference (2015)
7. Preston, C., Allan, V., Fraser-Thomas, J.: Facilitating positive youth development in elite youth hockey: Exploring coaches' capabilities, opportunities, and motivations. Journal of Applied Sport Psychology **33**(3), 302–320 (2021)
8. Preston, C., Fraser-Thomas, J.: Problematizing the pursuit of personal development and performance success: An autoethnography of a canadian elite youth ice hockey coach. The Sport Psychologist **32**(2), 102–113 (2018)
9. Schuckers, M., Brozowskia, L.: Referee analytics: An analysis of penalty rates by national hockey league officials. In: MIT Sloan Sports Analytics Conference (2012)

10. 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**, 1735–1757 (2017)
11. Wood, S.N.: Thin plate regression splines. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **65**(1), 95–114 (2003)
12. Wright, T., Trudel, P., Culver, D.: Learning how to coach: the different learning situations reported by youth ice hockey coaches. Physical Education and Sport Pedagogy **12**(2), 127–144 (2007)