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.

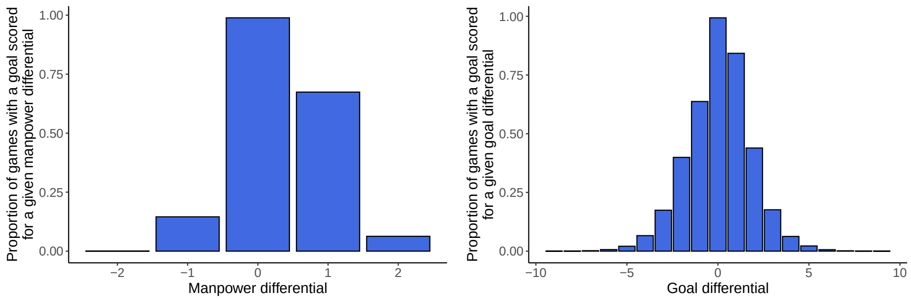

Fig. 1: Frequency of goals scored for different MD and GD.

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

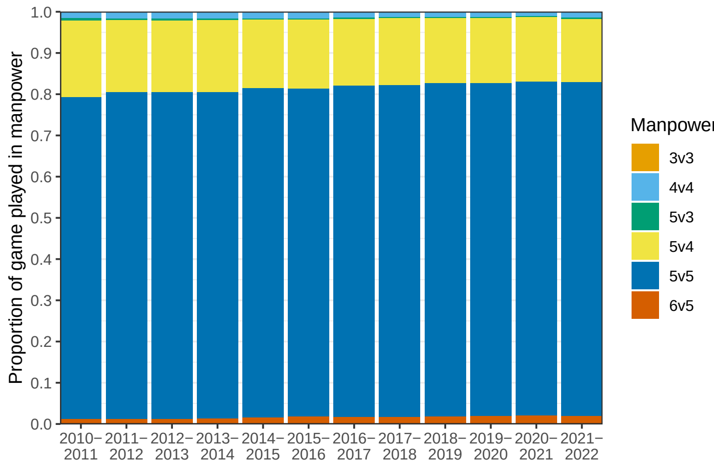

Fig. 2: Proportion of each manpower situation over time.

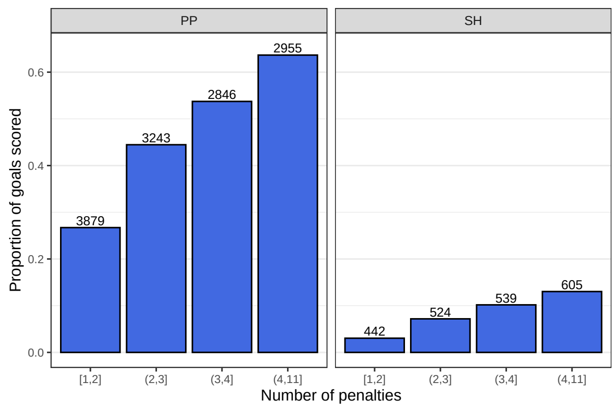

Fig. 3: Proportion of goals scored by number of penalties. The number on the bar shows the total number of goals.

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

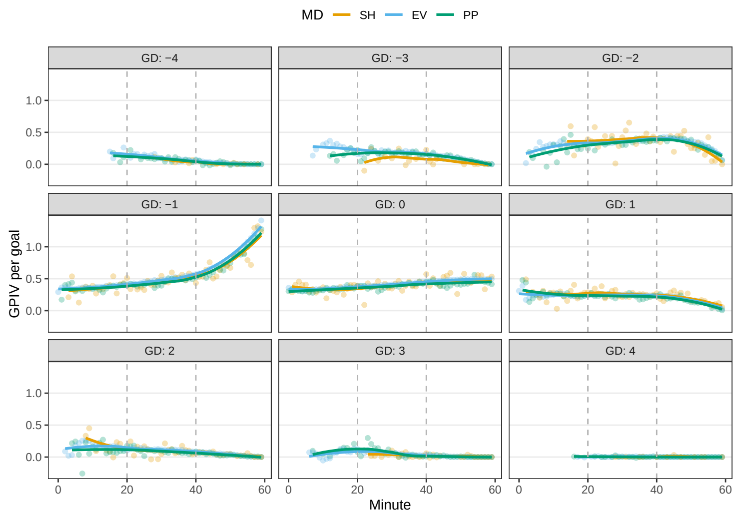

Fig. 4: GPIV per goal and minute.

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.
- SH_{te}: Interaction between shorthanded goals for and against per game.
- PP: Powerplay goal differential per game.
- PP_{te}: Interaction between powerplay goals for and against per game.
- SP: Powerplay and shorthanded goal differential per game.
- 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.
- 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.
- 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

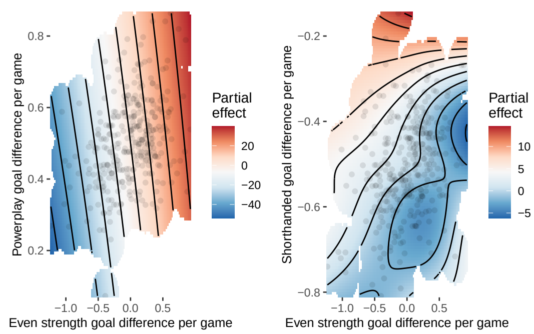

Fig. 6: Effect of manpower scenario goal differential on earning game points from the 'All' model. Partial effect of 0 means average, while red indicate more points and blue less points

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.

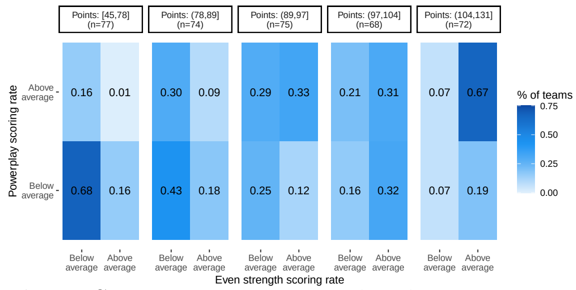

Fig. 7: Scoring rates compared to league average in PP and EV.

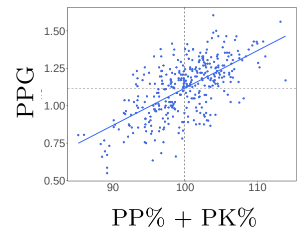

Fig. 8: Special teams (PP% + PK%) vs. points per game (all seasons).

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).

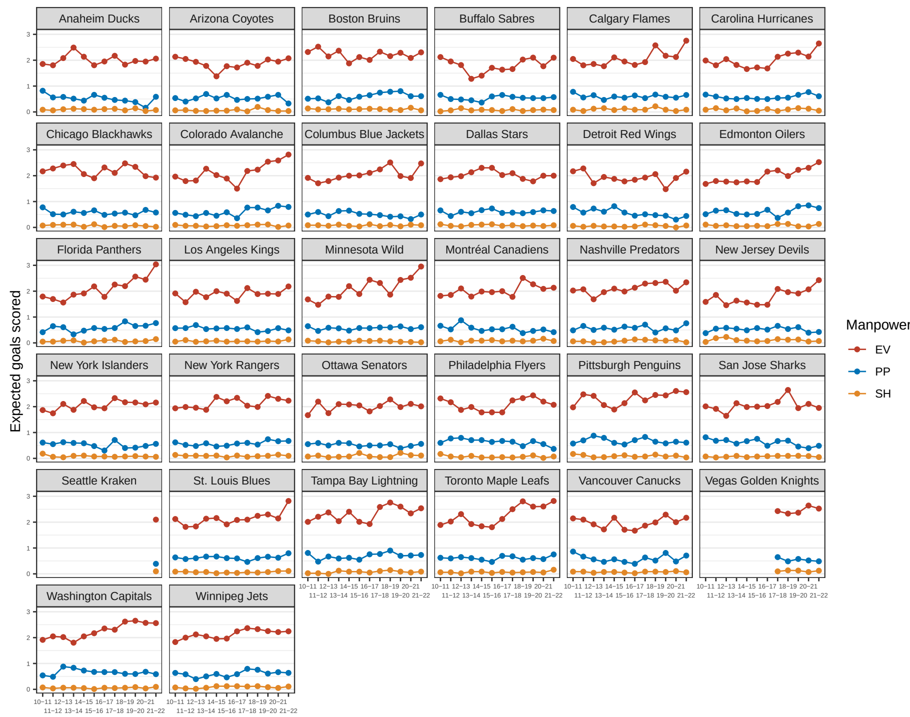

Fig. 9: Expected goals scored per team, season and manpower.

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)