Pilot Study on Optimizing Dump-In Strategy

Zachary Nasu

January 2020

1 Introduction

The purpose of this pilot study is to find if there exists a superior strategy to dumping the puck into the zone by way of puck placement. It can be seen from an article written by Eric Tulsky, Geoffrey Detweiler, Robert Spencer, and Corey Sznajder titled "Using Zone Entry Data To Separate Offensive, Neutral, And Defensive Zone Performance" [1] that offensive zone entries where the offensive team carries the puck into the zone generates almost twice as many shots as when the team dumps the puck into the zone. In addition, they also point out that trying to carry the puck in every time is also not an ideal as it would lead to a higher turnover rate. This being said, the idea behind the study was to find if there was a strategy that was ideal when dumping the puck into the offensive zone.

2 Methods

The data for this study was taken from a random sample of 36 periods from the 2018-2019 NHL season using the NHL source API *nhlsrcapi*. For the purpose of this study, a zone entry is categorized as a dump-in if:

- The puck is dumped-in from in between the opposing blue line and red line. (This includes tip-ins but does not include icings that are beat out)
- The puck breaks the plane of the bottom of the circle before any player regains control of the puck (Meant to omit area passes behind defenceman)
- The dump-in is not a shot on net

Through these conditions, there were over 600 data points that were collected through those 36 periods of NHL Hockey. Going further, these data points had the following variables:

1. Side of the ice
2. If a change occurred after the dump-in (Defined as if two or more forwards changed)
3. If the goalie stopped the puck behind the net or if the puck was picked up behind the net
4. If the team was able to recover the possession of the puck in-zone (either through gaining control for more than one second or from completing a pass)
5. Type of dump-in (Strong-Side, Rim, Cross)

There were a few assumptions made during data collection. One of those assumptions was that a goalie playing the puck behind the net is similar to a player picking the puck up behind the net. The reason why this assumption was made was that the defenseman could use the net as protection in either case and through experience, when a goalie plays a puck, the majority of the time the goalie just stops the puck for the defenseman, which has a similar effect to the puck being picked up behind the net without the goalie. In addition, another assumption that was made was that a dump-in should be considered a change if two or more forwards changed. This assumption was made as often when only forward changes, the impact on the forecheck is very limited as the new forward coming onto the ice is often the *high forward* or *F3*, which has a role to hang back and therefore, has a lot more time to get into the play. In addition, defenseman were not considered as they also do not play a big role in the initial forecheck when a puck is dumped into the

opposing zone. Going further, a case that came up occasionally was more than one forward changing early on the rush and the new forwards were able to catch up to the puck carrier before the puck was dumped-in. In this case, the dump-in was considered to have no change as, in theory, the change did not make an impact on the positioning of the forwards when the puck was dumped-in if they had not changed.

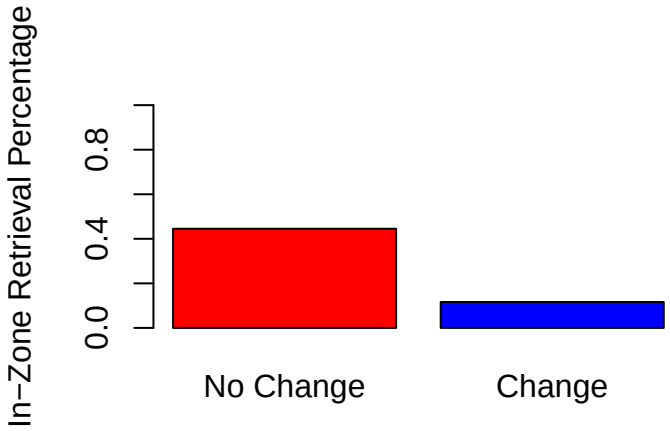
3 Findings

Through analysis, the variable that made the biggest difference on in-zone retrieval rates if the offensive team changed or not after the dump in. Tulsky, Detweiler, Spencer, and Sznajder[1] also made a similar observation, showing that the shots generated from a dump-in with a change were significantly lower than dumps without a change. These observations are connected as in hockey, possession is a necessity to create shots on goal. As seen below in Table 1 and Figure 1, the difference between the retrieval rates is quite significant and when we do a hypothesis test, we can reject the null hypothesis of $p_{Change} > p_{NoChange}$ at the $\alpha = 0.0001$ significant level.

Table 1: Retrieval Rates, Separated by Whether a Change Occurred

Type	Retrieval Rate
No Change	0.4451 ± 0.0447 (95% CI)
Change	0.1163 ± 0.0553 (95% CI)

Figure 1: Retrieval Rates, Separated by Whether a Change Occurred



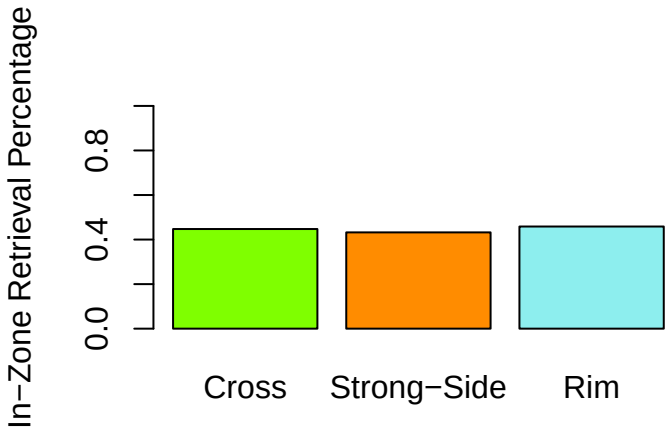
After analyzing these big differences, the two cases were then separated and analyzed further. Upon further analysis, the data where a change occurred could not yield accurate results as the data set was too small and further data collection will be needed. On the other hand, the data set where no change occurred was big enough to yield accurate results.

While isolating the impacts of different variables, unsurprisingly the side of the rinks seemed to have little to no effect on the retrieval rate. Conversely, what did surprise was how close the different types of dump-ins seemed to be and it can be seen in Table 2 and Figure 2.

Table 2: Retrieval Rates When No Change Occurred, Separated by Type

Type	Retrieval Rate
Cross Dump	0.4473± 0.1476 (95% CI)
Strong-Side Dump	0.4322± 0.0587 (95% CI)
Rim Dump	0.4586± 0.0779 (95% CI)

Figure 2: Retrieval Rates When No Change Occurred, Separated by Type

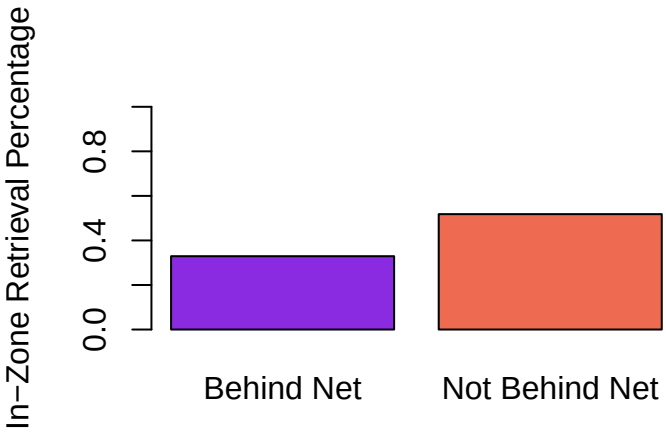


As it can be seen above, the different types of dump-ins look very similar, but since the confidence intervals are somewhat large, especially for the cross dump, it cannot be ruled out that there still may exist a superior method to dumping the puck in. Though a superior dump-in strategy was not found, there was one more interesting insight found in this study. Another big factor on recovery rates was if the goalie played the puck behind the net or if it was picked up behind the net and it can be seen in Figure 3 and Table 3 below.

Table 3: Retrieval Rates When No Change Occurred, Separated by if the Puck was Picked-Up Behind the Net

Type	Retrieval Rate
Behind Net	0.5081 ± 0.0559 (95% CI)
Not Behind Net	0.3293 ± 0.0712 (95% CI)

Figure 3: Retrieval Rates When No Change Occurred, Separated by if the Puck was Picked-Up Behind the Net

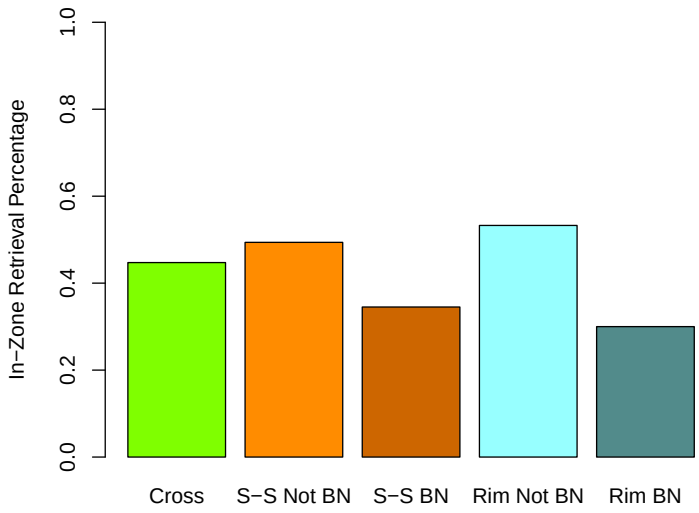


When doing a hypothesis test where we test if $p_{NotBehindNet} > p_{BehindNet}$, it yields an extremely low p-value and we can safely refute the null hypothesis at the $\alpha = 0.0005$ confidence level. Therefore, we can be fairly certain that having the goalie stop the puck or getting the puck stuck behind the net yields a lower in-zone retrieval rate. In addition, we can rework Figure 2 and Table 3 to include whether the puck was dumped in behind the net.

Table 4: Retrieval Rates When No Change Occurred, Separated by Type and If the Puck was Picked-Up Behind the Net

Type	Retrieval Rate
Cross	0.4473± 0.1476 (95% CI)
Strong-Side (Not Behind Net)	0.4938± 0.0775 (95% CI)
Strong-Side (Behind Net)	0.3451± 0.0877 (95% CI)
Rim (Not Behind Net)	0.5327± 0.0945 (95% CI)
Rim (Behind Net)	0.3000± 0.1270 (95% CI)

Figure 4: Retrieval Rates When No Change Occurred, Separated by Type and If the Puck was Picked-Up Behind the Net



For the cross dump, an accurate retrieval rate could not be obtained for when the puck goes behind the net as it is an extremely rate event. Building off of the previous hypothesis test on if the retrieval rate is higher when the puck is not stopped behind the net, further hypothesis tests can be done on both the Strong-Side and Rim dump. For the first hypothesis test, with the alternative hypothesis being $p_{S-S(\neg BN)} - p_{S-S(BN)} > 0$, we get a p-value of 0.0073. In addition, for the second hypothesis test where the alternative hypothesis is $p_{Rim(\neg BN)} - p_{Rim(BN)} > 0$ yields a p-value of 0.0032. With both of these p-values, we can conclude that there is very strong evidence against both null hypothesis', which are when the puck goes behind the net, there is a higher in-zone retrieval rate. Also, with these values, both null hypothesis' can be refuted at the $\alpha = 0.01$ confidence level.

4 Conclusion

Through the findings of this pilot study, it can be concluded with a high amount of certainty that one of the best dump-in strategies is to keep the puck from stopping behind the net. One of the reasons why this strategy is tangible is that most situations where the puck ends up behind the net, it can be prevented by good puck placement by the player who dumped the puck into the offensive zone. This strategy is not new to the game though and is the reason why goalies play the puck behind the net. The drop-in recovery rates when the puck stops behind the net is consistent with the logic that when a goalie gets to the puck first, they have control and can make a play to their own players and when the puck naturally stops behind the net, the defensemen have an advantage as they can use the net as a pick. Conversely, when the puck does not go behind the net, the in-zone recovery rates of around 50% follow the logic that it is a 50/50 puck battle when the puck is in the corner or along the half-wall in the case of a rim.

In addition, another factor that was found to have a big effect on in-zone recovery rates on whether the players changed or not after the puck was dumped in. This finding is not as helpful though in the way of strategy as line changes on the fly is a necessity and dumping the puck is in logically the best way to go about it.

Last but not least, the final finding of this case study, which was also the main purpose of the study, was to find if there was a dump-in type (Cross, Strong-Side, Rim) that yielded a higher recovery rate than the others, and from the data collected it is still inconclusive. From inspection of the data, all three types have very similar recovery rates, but since the sample is relatively small, there is a lot of uncertainty and there is

still the possibility that a superior type can still exist.

5 Sources of Error

1. Possible variability when recording the dump-in type variable. Generally a cross dump is fairly explicit, but occasionally the rim and strong-side dump is very similar. To combat this, those outliers were not recorded, but some might have been recorded anyway.

2. Human error. There might have been mistakes made when recording the data.

3. Possible error when distinguishing between a puck that was picked up behind the net or not. Due to the fact that they are is continuous and not discrete, there was some data points that were close and not explicitly one or the other and those for the most part were taken out.

6 Revision and Expansion

1. Record shots on net and shot attempts after a whistle if the play had not left the zone yet and stays in-zone after the whistle. The mistake made during this study was only counting the shots and shot attempts up to the whistle.

2. Differentiate if a goalie plays the puck or if the puck naturally behind the net. A study can be done on whether goalies should play the puck or not. The study can also include rate at which goalies make a mistake and cost goal. In addition, the difference in when they play the puck and don't play the puck can be used as goals saved when they play the puck.

3. Evaluate the cost of a turnover at the opposing blue line in terms of expected goals

4. Approximate a best ratio for dump in plays to carry in considering turnover rate, cost of turnover in terms of expected goals, expected goals for a carry in, expected goals for a dump, etc.

5. When a puck is stopped behind the net, add a variable that shows the goalie's hand. On a rim dump, the goalie might have a harder time stopping the puck on their backhand and could lead to strategic advantage.

References

[1] R. S. C. S. Eric Tulsky, Geoffrey Detweiler, *Using Zone Entry Data To Separate Offensive, Neutral, And Defensive Zone Performance*. MIT Sloan, 2013.