Universiteit Leiden
The Netherlands

Master Thesis

The ultimate way to play ultimate: Exploring performance indicators in European ultimate frisbee games.

---

by

Emma Boom
(s2326183)

First supervisor: Dr. Maarten H. Lamers
Second supervisor: Dr. Tiago G. Russomanno

July 17, 2025

The ultimate way to play ultimate: Exploring performance indicators in European ultimate frisbee games.

Emma Boom
Department of LIACS, Msc Media Technology, Leiden University
Leiden, The Netherlands
e.e.boom@umail.leidenuniv.nl

ABSTRACT

Context. Ultimate frisbee is a niche sport and quite unknown to the academic world. Even though sports performance analysis has been done in many popular sports such as football and basketball, it has been applied very little to ultimate.
Goal. Examining what performance indicators are the most predictive of success in European ultimate frisbee games.
Method. A set of metrics was designed, followed by the annotation of videos in which the data for these indicators were extracted from 16 recorded ultimate frisbee games from the EUC 2023. The last step was performing analyses to determine what predictors are predictive of success in European ultimate frisbee. Next to that, state transition probabilities were extracted from annotated field locations to gain insight into passing behaviour.
Results. From the paired-sample t-tests and points win margin analyses, attempted scores, score efficiency, disc possession, passing accuracy, total turnovers, number of substitutions, possession efficiency and passing rate were found to be significant. The state transition probabilities showed that winning teams kept their passes more within zones than losing teams.
Conclusions. It can be concluded that winning teams consistently demonstrate assertiveness on offense, control-oriented, risk-avoident and patient play, strong defense and good spatial awareness. Moreover, winning teams reveal better throwing connections and active defense than losing teams. State transition probabilities suggest that losing teams make longer throws, while winning teams keep the disc within and around a zone and advance slowly but steady.

Keywords: ultimate, ultimate frisbee, performance analysis, performance indicators, state transition probability.

1 INTRODUCTION

Most of the times when I tell people I play ultimate frisbee, they laugh and say they have done that too on the beach or in a park. Little do they know that ultimate frisbee goes a bit further than just simply throwing a disc back and forth to each other. It is a recognised team sport with a set of rules and practices, just like other team sports. And even though there is no "winning at all costs" mentality allowed in this sport (as can be seen in the spirit of the game aspect, see next section for an explanation), winning is still very important.

But why should we research this sport? We all want to win in sports, and some teams know how to do this better than others. Oftentimes, these teams gain their knowledge from years of experience, and teams can also include insights from performance analyses of their games into their strategies. Performance analyses can give a glimpse into the winning factors, also called performance indicators of a game. These can then be optimalised for success.

This has been done in many sports, but it is most apparent in popular sports such as football and basketball, where a lot of money is involved and the stakes are high. For ultimate frisbee this has been done once before, in a study by Lam et al. in 2021. They looked at recorded ultimate frisbee games and performed analyses post-game. They found some significant performance indicators, but their initial list of possible predictors was quite small and they looked at games from the American Ultimate Disc League (AUDL, now UFA), which is a different variant of the ultimate that is played in Europe (and mostly globally) and uses alternative rules and practices. The knowledge gap that this creates makes it fruitful to look into performance indicators in European ultimate.

As a solution approach, in this study we will look at recorded games from the European Ultimate Championships of 2023 (EUC 2023). From previous research and given the characterics of the sport, a list of possible predictors will be proposed. Through the annotation of the possible performance indicators of this list a dataset will be extracted, and analyses will be performed to gain insight in successful strategies. On top of that, passing behaviour will be analysed from the field transition probabilities.

First, an explanation of the sport and its rules will be laid out, followed by an extensive outline of performance analysis in team invasion sports, common performance indicators in field based team sports, ultimate frisbee characteristics compared to the popular and widely researched sports, and an in-depth discussion of existing performance analysis in ultimate frisbee. Next, the aim of the research is stated. Then, the methods and results are split into three parts: the list of possible performance indicators, the annotation of the games, and the analyses of the indicators. Afterwards, the study will be thoroughly discussed and drawn to a close with a conclusion.

2 ULTIMATE FRISBEE

If you are unfamiliar with ultimate frisbee, the following section will bring you up to speed on the sport. Ultimate frisbee (or simply called "ultimate") is a team invasion sport, which defines itself as a team game characterised by the objective of invading the opponent's territory to score points while simultaneously minimizing the opposing team's score, all within a specified time limit (Gréhaigne et al., 1999). Invasion sports are played with a certain object, which is usually, but not limited to, a ball. In ultimate frisbee a disc is used. Points can be scored by moving the object into the target area,

which can be a net or a specific zone. Maintaining possession of the object and progressing towards the oppposition's defensive zone are offensive tactics, while preventing or slowing movement of the object towards the defensive area, protecting the target area and trying to regain possession of the object are defensive tactics.

Ultimate frisbee is a relatively new sport, developed in 1968 at a parking lot of Columbia High School. Since then, it has had a slow but steady rise in popularity, being most popular in university sports settings, but also being picked up by children, amateurs and professionally. There are three aspects that makes the sport stand out from other sports: it is self-refereed, a non-contact sport and pays close attention to sportmanship or "spirit of the game". This spirit of the game aspect is assessed after each game by both teams. First, right after finishing the game, each team gets together and discusses points the other team could improve on regarding "spirit", and shares these points in a "spirit circle" where players of both teams are standing alternately in a circle. Afterwards, each team individually rates the other team on five markers (1. rules, knowledge and use, 2. avoiding body contact, 3. fair-mindedness, 4. attitude and self control, and 5. communication). Thus, in a competition, there is a game winner and a spirit winner. Moreover, it is a mixed gender sports with an open, women and mixed division. There is no specific men division, but in the highest level of the open division it is not uncommon for it to be fully male. Having said so, in a game there are female matching players (FMP) and male matching players (MMP). In the mixed division, a 3/4 ratio of FMP and MMP are in play, changing according to a specific rhythm (usually "ABBA"). On ground of locations, ultimate frisbee can be played indoor, outdoor and on the beach, with each playing ground having a variation of the rules to accommodate the environment (see Figures 1 and 2).

Alright, this ultimate crash course is all very interesting, but how do you play it? The rules are based off and similar to sports like football, basketball and rugby. In short, according to the World Flying Disc Federation (WFDF) rules, the goal of the game is to score points by catching the disc in the opposition's end zone (World Flying Disc Federation, 2025). The game is played on a 100 meter by 37 meter field (that is for outdoor, for indoor and beach the field is smaller), with an end zone at each outer edge. For visualisation: it is about as long as a football field, but twice as narrow (see Figure 3). There are two opposing teams, each starting at the beginning of a point at their defending end zone. To start the point, one team performs a "pull", where the disc is thrown towards the opposition's zone as far as possible. This is where the game begins, with the disc being caught or picked up by the opposing team. If the disc from the pull lands outside of the field, it can be "bricked", and this leads the disc to be placed at the the point in the middle of the field on the offensive team's side, called the "brick mark", from where the game will be continued (see Figure 3). From there on, the disc has to be passed to team players in any direction, but the thrower cannot move while holding the disc. Offensive players make sprints in the field, called "cuts", to receive the disc. While in flight, the disc can be intercepted by the other team. If the disc is not caught by a player of the same team (i.e., by touching the ground), the other team becomes offense. Each player has a certain amount of time to throw the disc, counted (referred to as "stalling") by their defender. Once a disc has been caught in the opposition's end zone, a point is made. After each point, teams can substitute as many players as they want from their team.

Figure 1: Ultimate frisbee on a grass field.
Image courtesy of kreatif.minds.

Sometimes, game violations occur, and these conflicts are resolved among the players themselves: first by those involved, but assistance can be requested from other players or the sideline. To communicate to all players and spectators what is being discussed, specific hand signals are used and repeated by the other players. This makes communication an important part of the self-refereeing. A violation can be either accepted or contested, and the game continues by the disc going back a pass or continuing from there.

Now that the rules for playing an ultimate frisbee game are clear, we will name or repeat certain terms that are important to understand for this thesis. Logically, points can be scored, and breakpoints are points that are scored by the team that started the point on defense. This is seen as very positive, since the team had a disadvantage at the beginning of the point and nevertheless managed to score. A turnover is the loss of possession of the disc to the opposing team. A pull is the first throw at the beginning of the point, thrown by the defensive team to the offensive team, and can be bricked when it was thrown out of bounds. A stall-out means that the offensive player with the disc has taken too long to throw, and a turnover is in place. With all this in mind, now you are ready to embark on this thesis.

3 RELATED WORK

Before diving into the specifics of the study, there are first some important terms and studies that the reader needs to know to get on the same page concerning performance analysis in ultimate frisbee.

3.1 Performance analysis in team invasion sports

As mentioned in the previous section, team invasion sports are team games where the main objective is to invade the opponent's territory while defending your own. Tactical behavior is a crucial factor influencing performance in team sports such as ultimate frisbee. It is important for a team to have multiple strategies to manage spatial positioning to achieve a common objective (i.e., scoring points).

Figure 2: Beach ultimate.
Image courtesy of Sam Kehl.

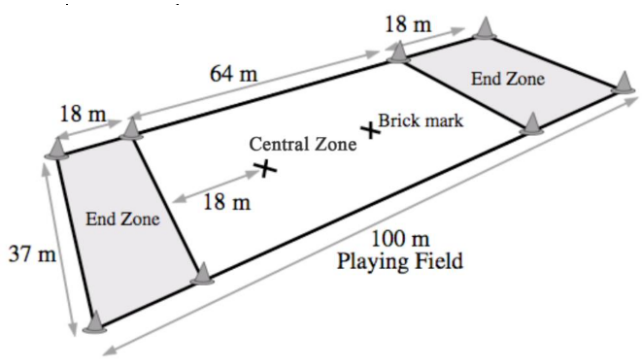

Figure 3: Ultimate frisbee field dimensions.
Source: WFDF Rules of Ultimate 2025–2028, produced by the WFDF Ultimate Rules Sub-Committee.

Simultaneously, knowing how to respond to the opponent's actions within the constraints of the game environment is a crucial part of a team's strategy. Strategies are playing plans discussed in advance (both offensive and defensive) intended to be applied during a game, while tactics are playing plans that are developed during a game in response to the opponent's play (Gréhaigne et al., 1999).

Performance analysis is to look at the results of strategies and tactics and analyse this data to detect factors that are related to success. These factors are called performance indicators (Hughes and Bartlett, 2002). The more data is collected, the more patterns can be detected. In team-based sports, success means scoring more points than your direct opponent. The review article by Lord et al. (2020) lays a groundwork for analysing performance indicators in invasion sports. From 537 articles on performance indicators in invasion sports, they identified six themes of performance indicators. For each theme, they looked at the insights that could be gained and how they relate to team performance.

The first theme they identify is game actions, which are game events measured per team as a frequency or percentage (think of percentage of possession per game). Closely related is dynamic game actions, the second theme, where game actions are related to the spatial location on the playing field where it occurred (i.e., possession per area). Movement patterns indicate a succession of actions from possession to a certain outcome, which they identify as a third theme. An example of a movement pattern is corner routines, such as can be seen in socccer and field hockey. As a fourth theme, they discuss collective team behaviour, which looks at the field surface area that is utilised during a game. The fifth theme is social network analysis, which identifies key players through team passing networks. The last theme they identify is game styles, in which the strategies used by teams is analysed through linking dynamic game actions to speed of play, which are then clustered into attacking and defensive styles.

Even though the approach of simply looking at game actions has been applied most often in the field of sport performance analysis, the authors argue against it. Game actions are outcome-oriented variables, which informs about what happens in a game, but does not explain how or why such an event happened, and how to recreate such an event (e.g., a goal-scoring opportunity in football). They expose what indicators are and are not related to succes, and as such it can be derived that to succeed, a team has to be better in those indicators than the opponent. To gain insight into the context of an outcome-oriented variable, using dynamic game actions could be better, because including space and opposition interaction in game actions provides details about how a team functions in different game situations. Movement patterns provide process-oriented variables, which incorporates space, time and player-opponent interaction to paint a broader picture of all features that can influence game outcome and performance. Overall, the authors emphasise looking at games in a holistic manner: seeing game performance as a whole and including relations between events instead of looking at individual events on itself. This allows the dynamic and complex interactions that emerge in team invasion sports to be captured. Morgulev and Lebed (2024) also recommend to look further than just game actions. In their article, they critique research where analyses on performance indicators are shallow or close replications of previous studies without taking into account the dynamic aspect of team sports (i.e., only looking at game actions and not the whole picture). Like Lord et al. (2020), they also suggest to include time, space and interaction when annotating game actions. Moreover, it is recommended to normalise performance indicators (e.g., number of turnovers divided by the number of possessions) (Lam et al., 2021).

3.2 Common performance indicators in field based team sports

Where there is money, there is research. Therefore, in popular sports like football, basketball and rugby a lot of research is done in terms of performance analysis (Ávila Moreno et al., 2018; Lord et al., 2020). Even though these sports are not ultimate frisbee, they share some similarities in their use of field and playing mechanics. Therefore, the results from these studies will be used as a base for this research, but they will be used with caution. What is also noticeable in big sports, is that there is much more data available online with statistics that can be freely requested (Plakias et al., 2024; Woods et al., 2017). Performing analyses with large datasets therefore makes it easier for these sports.

3.2.1 Performance analysis in football. From the review study on performance indicators in general by Lord et al. (2020), it is clear that most research on performance analysis has been done in football. Many papers and, more importantly, review papers and meta analyses can be found on this topic revealing important insights into game predictors.

For example, a large review study has been done by Lepschy et al. (2018), where they identified performance indicators and methods from 68 articles. From the 76 performance indicators that they found, the most relevant predictors that came out of their review are efficiency (i.e., ratio of goals over number of shots), shots on goal, ball possession (measured in time of possession), passing accuracy (percentage of successful passes over total passes), quality of opponent and match location. Most of the articles collected their data through secondary data (i.e., freely available data from sports organiser websites) or (mostly manual) video analysis. They emphasise the importance of a large enough sample size, since many of the articles lacked that. Thus, looking at a whole season or competition is a good practice.

Thus, from performance analysis studies in football, performance indicators that were found are efficiency, shots on goal, ball possession and passing accuracy. Even though they are relevant performance indicators, there is an important difference between football and ultimate frisbee. That is: football is a low-scoring game sport where scoring a goal is a rare event. In sports such as football, scoring a "lucky" point can have a large impact on the outcome, while in high scoring games a lucky point may not make much of a difference. For example, in football it has been found that scoring the first point can have a large influence on winning (Lago-Peñas et al., 2016), an effect that may not be present in high-scoring game sports. However, since football is the sport with the largest database on performance analyses, it can be fruitful to look at review studies. Nevertheless, looking at high scoring game sports such as basketball or rugby can reveal other important insights.

3.2.2 Performance analysis in basketball. After football, performance analysis in basketball is also a hot topic according to the review study by Lord et al. (2020). On top of that, basketball is a high scoring game, therefore a deeper apprehension might be gained from looking at performance research in this sport.

In a study by Çene (2018), the author looked at data acquired from the Euroleague website that collects boxscores for each game. Boxscores are the results of a game and player performance statistics. Due to the results already being available, they were able to include 259 games in their analyses. They did a k-means cluster analysis to divide groups based on final score differences into close games (<10 points difference), balanced games (between 10 and 21 points difference) and unbalanced games (>21 points difference). Next, they executed different statistics to understand the pattern on game statistics between winning and losing teams and found that different performance indicators are relevant for each group. In close games, true shooting percentage (i.e., quality of the shot instead of quantity), steals (i.e., gaining possession of the ball from the opposing team) and committed fouls (i.e., total violations of the rules) were found to be of importance, while in balanced games 2- and 3-point field goals, steals and defensive rebounds (i.e., the player that catches the ball first after the ball jumps out of the basket due to successful defense) were found. Defensive rebounds and 2-point field goals were most important in unbalanced games. Overall, it can be concluded that both direct attacking behaviour (point field goals) and defensive behaviour (steals and defensive rebounds) are most influential for winning a game.

Another study looked at over 328 matches and 50 teams (Plakias et al., 2024). In their study, they found that the most significant performance indicators are effectiveness of field goals (i.e., whether they scored a point from a 3-point position), assists-to-turnovers (i.e., how many assists and turnovers a specific player has had), rebounds (i.e., after the ball jumps out of the basket, the player that catches the ball first), defensive ratings (i.e., percentage of number of points scored by a player per 100 team possessions), and steals.

Another study looked at 156 female basketball matches from 4 consecutive Olympic games (2004-2016), with data acquired from the official Olympic websites (Leicht et al., 2017). They found that shooting proficiency (percentage of goals made over goals attempted) and defensive actions (such as defensive rebounds, steals and turnovers) were the best performance indicators.

In all three discussed basketball performance studies, shooting proficiency (which is similar to efficiency in the football studies) and defensive actions were found to be most related to winning. However, these studies only looked at static variables and did not include time or space in their analyses. Moreover, a basketball field is quite small compared to an ultimate frisbee field. Different dynamics may therefore be at play and sports with similar field sizes could be of interest to look into.

3.2.3 Performance analysis in rugby. Less popular, but still towering high above other sports performance studies is the sport of rugby and its relatives American, Australian and Gaelic Football (Lord et al., 2020). Sports from the rugby family may even be most related to ultimate frisbee. It is therefore wise to include performance analyses into our search for performance indicators.

In a study by Woods et al. (2017), they gathered 376 observations from rugby games in the 2016 National Rugby League in Australia and New Zealand. They looked at wins/losses and ladder position (1 to 16, 1 being best). The indicators they found that explained match outcome and ladder position the best included try assists (i.e., assisting a score by bringing the ball into the opponents' end zone), line breaks (i.e., breaking the defensive line of the opposition), all run metres (i.e., total metres run forwards on offensive), missed tackles (i.e., line breaking without being tackled), kick metres (i.e., how far the ball travels with each kick, accumulated), dummy half runs (i.e., number of times the offensive team runs before passing the ball) and offloads (i.e., passing the ball to a teammate whilst being tackled). All in all, offensive actions were found to be most predictive of winning.

Looking at rugby studies, offensive actions are found to be most predictive of winning. However, offensive behaviour is regarded mostly in rugby-specific predictors, which can be difficult to translate to other sports.

In many sports studies, winning has been associated with offense tactics (Mikolajec et al., 2013; Woods et al., 2017) whilst in other studies defensive actions were found to be most indicative of success (Leicht et al., 2017; Plakias et al., 2024). This really depends on the type of sport.

Although these studies examined indicators of success in football, basketball and rugby, some limitations and/or methodological problems in these studies can be observed, which leaves room for improvement and the suggestion for predictors specific to the ultimate frisbee sport. The main limitation is that these studies are based on fundamentally different sports. In rugby and basketball studies, there are multiple ways to score (e.g., scoring through two poles, catching in end zone, scoring from different locations). These indicators would all have to be merged together into one indicator for the current study. Moreover, some of these sports, such as basketball, games are played with quarters or halves, and the breaks can be quite long. In ultimate frisbee, breaks are halfway based on time or points, and the break is kept to a maximum of five minutes. Looking at specific ultimate frisbee characteristics can provide a clear picture of what factors play a role in an ultimate frisbee game.

3.3 Ultimate frisbee characteristics

As ultimate frisbee is a team invasion sport, many similarities are found, but distinct differences differentiate the sport from other team invasion sports. First of all, being a team invasion sport, all sports that are performed individually (e.g., golf or tennis) or without a (single) direct opponent (e.g., track and field) are fundamentally different to such an extent that the performance indicators from those sports cannot be used for the sport of ultimate frisbee. However, in tennis, it is possible to play in a team, but you cannot pass the ball to your own teammate, which is the only useful way of passing the disc in ultimate. Then, recurrent performance indicators in team invasion sports mention passing behaviour. Passing behaviour in ultimate frisbee is solely to your own team, you should only pass to your own teammates, and not to the opponent eventually (as in volleyball). Moreover, you score by catching the disc in the opponent's end zone, similar to rugby or American football.

As explained in the introduction, ultimate frisbee has some aspects that are unique to the sport. Even though the sport is team based, no contact between opponents is allowed on the field, unless agreed upon and kept to the minimum. At no time is a dedicated referee present, as players and the side line are expected to resolve conflicts during the game. This requires a level of self-awareness at all times, and can also use up valuable playing time, even though conversation is kept as short and as clearly communicated to other players as possible. In all sports mentioned before, possession of the ball is indefinitely. Even though in basketball there is a time limit to how long you can hold the ball without dribbling, in ultimate frisbee the player has to throw the disc within 10 seconds, with the stall counted aloud by their defender. This time pressure forces quick thinking, influencing team strategies. Another unique aspect is the freedom to switch substitute players after every point, and when a player gets injured. Especially in higher leagues, teams make sure to have at least two full lines of players to alternate playing points. This allows for players to get more rest in between points, but also makes the game more dynamic with more different players and playing styles. Also, this makes it really a team effort instead of a single person effort.

In widely known sports, there are already many metrics freely available online. For example, for basketball there is the Euroleague website (www.euroleague.net), where useful metrics for each basketball match and player are recorded. Many studies make use of these metrics to perform their analyses, making it easier to include large datasets for their analyses. For ultimate frisbee, only some large competitions or American leagues have data recorded, and only since the last couple of years. This can be found on the official European Ultimate Federation (EUF) website that records statistics for large events, but only data for scored points, assists and its timestamps are recorded (https://eucs-schedule.ultimatefederation.eu). For this study, more data than provided on the EUF website is preferred.

3.4 Performance indicators in ultimate frisbee

So far, there has been one other study on tactical performance analysis in ultimate frisbee (Lam et al., 2021). In this regard, their study offers a framework for analysing certain performance indicators in ultimate frisbee. In their study, they investigated the spatial structure of possessions and passing patterns. For this, they adapted existing methods of state transition modelling to suit the sport of ultimate frisbee. To annotate performance indicators from 14 games of the American Ultimate Disc League (AUDL) from 2017, an observational system was designed using Microsoft Excel.

In total, they had identified nine performance indicators which they split up into three categories: general match indicators, technical, and tactical behaviours in ultimate frisbee (see Table 1). They looked at point and break points scored per game, possession opportunity (i.e., percentage of starting on offense for both teams in a game), total (of which forced and unforced) turnovers per game, completed passes per game, average number of passes per possession and turnover-to-point conversion efficiency (i.e., how often a team that has won disc possession scores a point compared to total of opponents turnovers).

Their adapted state transition model accounts for disc possession and loss in certain parts of the playing field, and they map all the possible passing patterns. Then, through analysing freely online available video footage, they annotated 14 games and performed descriptive analyses on the data. Points and breakpoints per game and turnovers (total, forced and unforced) were found to be significant performance indicators. Moreover, they calculated transition probabilities, i.e., how likely a pass or score would have been made from one part of the field to another part of the field, or where certain actions (such as turnovers) were most often committed.

From their results, they conclude that most of the action happens in the middle zone, which they attribute to the fact that the middle zone is larger than any of the other zones. Moreover, winning teams committed more forced turnovers in the middle zone than losing teams, and endured fewer turnovers overall. The authors argue that winning teams conserve better disc possession and they suspect that these teams could be better by making fewer mistakes (as they commit fewer unforced turnovers) and have a better defence on the other team (as they have fewer forced turnovers imposed on them).

Nevertheless, the authors identified only nine performance indicators, which is quite low compared to other sports performance studies. Of these nine indicators, five were found to be significant, of which one is the amount of points scored which is logical and does not really give additional information. One of the limitations

that they mention is that they only looked at 14 games, and they thus recommend to use a larger data base.

In addition to these limitations, they base their rules on AUDL rules, which are different from the official WFDF rules which are used in Europe. For example, in the AUDL games there is a referee present, while in common ultimate the games are self-refereed. The stall count in AUDL is slightly shorter than in WFDF games (7 vs 10 seconds), which is counted by referees instead of the marker. Moreover, AUDL games consist of four quarters of 12 minutes, while WFDF games consist of two halves of 35 minutes. In AUDL, the playing field is larger while the end zones are shorter, and fouls are penalised with backward movement of the disc. There is also no maximum amount of points; instead there is a maximum playing time which only clocks while the disc is in play. In WFDF games, a game ends when the maximum amount of points is reached (usually 15 points) or when the time is up (which depends per competition). In AUDL games, teams are allowed to change substitute players at any time except for after scores and timeouts, while in WFDF games only after points and injury calls. AUDL games are in general very commercial, whilst in WFDF games there is a stronger focus on the "spirit of the game" with traditional rules. Overall, the differences that are found between the AUDL games and WFDF games are similar to those found between ultimate and other sports. Moreover, the larger field and continuous games could impact gameplay strategies and performance metrics. It is therefore useful to treat our current study as a separate performance analysis on ultimate frisbee.

3.5 Key findings

By looking at previous research, we learn the following things:

- It is important to include time and space in the analyses by looking at dynamic game actions, and not simply at the actions in itself. Linking dynamic game actions to speed of play through field locations can be used to identify game styles.
- Normalising indicators makes them more widely applicable.
- The most important performance indicators found in previous studies in sports that are similar and relevant to ultimate frisbee are efficiency (or shooting proficiency), shots on goal, ball possession, passing accuracy, assist-to-turnovers, and (forced) turnovers. However, certain relevant performance indicators were so specific to the sport that they are not fitting for this study, leaving room for exploration of ultimate-specific indicators. In addition, assist-to-turnovers is a player-specific performance indicator, and the rest are team-specific performance indicators.
- In ultimate frisbee, performance analysis has only been done in 2021 by Lam et al. in combination with State Transition Modeling, in which they found that points and breakpoints per game and turnovers (total, forced and unforced) were significant performance indicators. However, their initial list of identified performance indicators is quite short and they looked at teams that used game rules from the AUDL, which are to a slight extent different from the rules used in European ultimate, leaving room for more exploration of performance indicators.

4 RESEARCH STATEMENT

Now that all the terminology, rules and previous studies are cleared up, we will focus on the study at hand. The aim of the current study is to examine performance indicators in European ultimate frisbee games, i.e., the factors that are not only related to success, but are also the most predictive of success. The main research question is therefore: *What performance indicators are the most predictive of success in European ultimate frisbee games?*

To answer this question, this study will be divided into three parts: designing metrics, extracting data from games and what makes a winner. For designing metrics, performance indicators found in earlier research and a proposal for novel performance indicators will be presented. This yields a set of metrics (i.e., possible performance indicators). In the extracting data from games part, the set of metrics will be extracted from recorded elite games for a post-game analysis. The annotation of these performance indicators will yield a dataset. The last part, what makes a winner, is where statistical analyses will be performed on the dataset. This yields a list of what performance indicators are the best predictors of winning ultimate frisbee games, to optimalise winning.

5 DESIGNING METRICS

As mentioned before, the current work attempts to identify significant performance indicators in European ultimate frisbee that have been found in other sports, and attempts to investigate more possible indicators.

5.1 Performance indicators from previous studies

Based on previous and related studies, certain indicators such as efficiency (or shooting proficiency), shots on goal, ball possession, passing accuracy, assist-to-turnovers and (forced) turnovers were found to have a good predictive ability of winning sports games. Since for this study we will look into team-specific predictors, assist-to-turnovers are not fit to look at for this study. We will thus try to translate the remaining indicators to fit the ultimate frisbee environment to see if they are also predictive of winning ultimate frisbee games, and we will normalise them whenever we can. For example, in the first performance indicator in the above list (efficiency, which is ratio of goals over number of shots), goals will be scores and shots will be attempted scores, and ball possession will be disc possession.

From the ultimate frisbee research by Lam et al. (2021), points, breakpoints, total turnover, forced turnovers and unforced turnovers were found to be significant. Since points and breakpoints are logic, we will omit those but investigate the three other indicators (total, forced and unforced turnovers) in the research to see if they are also predictive of success in European ultimate frisbee.

We expect that the interpretation on whether a turnover is forced or unforced can be difficult to determine in some situations. This makes it hard to objectively annotate these indicators. If it turns out that one or both turnover options play a major role in prediction of success, then we can assume the interpretation has gone properly. If it turns out that forced or unforced turnovers do not play a role, then there are two explanatory possibilities: 1) forced and unforced

Table 1: *Performance indicators used in the Lam et al. (2021) study.*

Category	*Performance indicator*	*Description*
General match indicators	Points scored per game*	Points scored in a game.
	Break points scored per game*	Points scored by the team that did not start on offense.
	Possession opportunity	Percentage of starting on offense for both teams in a game.
Technical behaviours	Total turnovers per game*	How many times there was a loss of possession of the disc.
	Total forced turnovers per game*	Loss of possession due to active defense.
	Total unforced turnovers per game*	Loss of possession due to a mistake.
Tactical behaviours	Completed passes per game	A successful passage of the disc between two team members.
	Average number of passes per possession	How many times the disc was successfully passed between turnovers.
	Turnover-to-point conversion efficiency	How often a team that has won disc possession scores a point compared to total of opponents turnovers.

* indicators found to be significant in the Lam. et al (2021) study

turnovers truly do not play a role, or 2) something went wrong with the interpretation of forced and unforced turnovers.

5.2 Novel performance indicators

Besides the known predictors, we also propose some indicators that we suspect to have some influence on performance in ultimate frisbee games. As a first indicator, we propose number of subs (i.e., substitute players). Since more subs means more rest in between points, and thus more energy saved up for each point they play and less tiredness. On the other hand, more players could lead to more unexpectedness. A second proposed indicator is early wins; when the game is finished by time instead of maximum points. As a third indicator, we propose possession efficiency, which is the number of passes in each successful scoring attempt. A fourth indicator we propose is passing rate, which relates to the number of passes per minute on possession and which could indicate the team's playing style (e.g., fast and flowing or slow and controlled). As a fifth and last indicator, we propose possession duration, meaning the average time a team maintains possession before scoring or losing the disc.

5.3 Proposed metrics

Taken all these indicators together, the total list of possible performance indicators is attempted scores, score efficiency (percentage of scores over number of attempted scores), disc possession (percentage of time on offense over time on defense), passing accuracy (percentage of successful passes over total passes), turnovers (forced, unforced and total), number of subs, early wins, possession efficiency, passing rate, and possession duration. For an overview see Table 2.

6 EXTRACTING DATA FROM GAMES

From previous research, a good approach to identify performance indicators is to look at dynamic game actions, i.e., game events related to the spatial location on the playing field (Lord et al., 2020). For this study, with the means available (i.e., the recorded videos), investigating dynamic game actions and using those to identify game styles is the best course of action.

It is valuable that a domain expert interprets the game and circumstances, so an ultimate frisbee player with approximately four years of experience has performed the annotation.

6.1 Methods and materials

Ethical review was done by Leiden University and approved. No personally identifiable data was collected or analysed. All performance analyses were conducted at the team level using publicly available footage accessed via YouTube.

Therefore, in this study, a tactical analysis is performed, with data manually extracted post event from video footage of 16 ultimate frisbee games from the EUC 2023 to determine performance predictors in ultimate frisbee games. The EUC 2023 was chosen because 60 games are recorded and freely available on Youtube from the ulti.TV channel and additional statistical data can be found on the European Ultimate Federation (EUF) website (https://euc-schedule.ultimatefederation.eu). After personal communication with ulti.TV, they agreed on the use of their materials.

For this study, we chose to only look at elite teams. This was done to ensure that the athletes performed proper game execution, and so that mistakes were not made due to inexperience with the sport. Moreover, elite teams have good knowledge of the rules and plays of the game, and more advanced strategies could be employed than with lower level teams. On top of that, in elite teams the level of playing is similar between teams, whereas in amateur teams the differences between winning and losing teams can be exaggerated due to stronger teams playing against weaker teams. It is important that teams are of equal strength to be able to compare them and interpret performance indicators properly (O'Donoghue et al., 2008). Moreover, the teams are evenly matched in the number of games played. It is critical to examine indicators of game winning in elite teams to create a greater understanding of the aspects that contribute towards winning.

As discussed, in ultimate frisbee there are three divisions: open, women's and mixed. For this analysis, we chose to look at the mixed division because we think that the choice of division does not matter for the results and thus the personal preference was followed.

Table 2: *Proposed metrics in this study.*

Category	*Performance indicator*	*Description*
Previous research	1. Attempted scores[a]	Total scoring attempts.
	2. Score efficiency[a,b]	Percentage of points scored to total scoring attempts.
	3. Disc possession[a]	Percentage of time spent on offense to time spent on defense.
	4. Passing accuracy[a]	Percentage of completed passes to total pass attempts.
	5. Forced turnovers[c]	Turnovers caused by active defensive pressure.
	6. Unforced turnovers[c]	Turnovers due to execution errors or miscommunication.
	7. Total turnovers[c]	Total number of times a team lost disc possession.
Novel indicators	8. Number of subs	Number of substitute players used during a game.
	9. Early wins	Whether the game finished by maximum points instead of time.
	10. Possession efficiency	Number of passes in each succesful scoring attempt.
	11. Passing rate	Number of passes made per minute during possession.
	12. Possession duration	Time in seconds a team holds possession before scoring or losing the disc.

[a] found significant across several studies, [b] found significant across multiple sports, [c] found significant in Lam et al. (2021) study.

For the annotation, BORIS was used (Behavioral Observation Research Interactive Software, Friard and Gamba, 2016). BORIS uses the terms events and states, in which events are specific occurences of behaviours that happen at a point in time, and states are conditions that happen for a particular time with a start and stop indication. A project was created in which all the necessary events (such as points and passes) and states (such as which team is on offense) were coded to be able to gather the data needed to measure the indicators. The full games were downloaded from YouTube and opened in BORIS for annotation.

6.2 The annotation of games

First, one game had been annotated as a preliminary version to verify how doable the research is, and whether the proposed performance indicators are able to be extracted from the videos. A preliminary data analysis was therefore performed to extract the data and changes were made to the annotation method based on the results. In some situations, the course of action was hard to determine, so a table with annotation rules and definitions has been made to ensure that the same annotation is made in similar situations (see Table 3). The data of this first game will be used in further analyses. An overview of all the games can be found in Appendix A. The notes on the videos can be found in Appendix B.

For the annotation, we will base the division of the field on the study by Lam et al. (2021). Different from their study, we decided to divide the field in six zones instead of five, because in their study they observe that the middle zone was played on the most, but this was often attributed to the fact that that part of the field was also spatially the largest. In this research we want to find out if these specific middle zones have any influence on the game. Moreover, Lam et al. (2021) use a field where clear markings are visible for each zone and in our case this is not as evident, thus we will divide the field into six equivalent zones: two end zones (EZ_Off and EZ_Def), two red zones (close to the attacking and defending end zones: from the inner end zone line to the brick mark: RZ_Off and RZ_Def) and two mid zones (from the brick mark to the middle of the field: Mid_Off and Mid_Def) (see Figure 4). The notation of the location where the disc is caught will be relative to the team on offense, so if

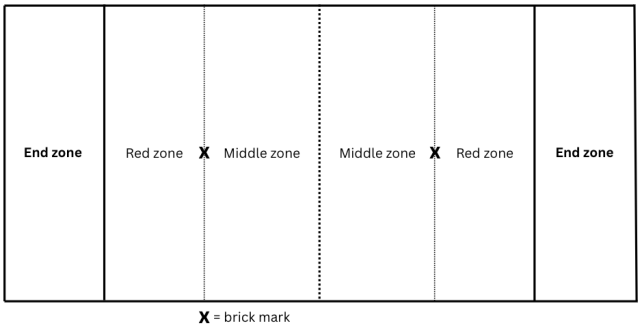

Figure 4: Field division in this study.

the disc is caught near the end zone where the team needs to score, this is annotated as RZ_Off. From these annotations, we will report state transition probabilities based off the ultimate frisbee adapted State Transition Model as described by Lam et al. (2021). Here, they define starting states (i.e., an event that initiates a new possession, such as the pull, or an earned turnover) and absorbing states (i.e., an event that terminates a new possession, such as a turnover or a score). Transient states are the different field locations. We name the states "Throw_From" and "Throw_To".

The boundaries of each spatial zone are sometimes hard to see because the footage of the games is not static; i.e., the camera moves with the disc. This could be of influence on the interpretation of where a disc was caught.

Each game yields a raw dataset with annotated events. A Python script was written to clean the data so that it can be used for the analyses. The whole annotation process provided us with a raw events table for each game which we processed into a summary indicators table and an game stats table. The csv files of the dataset, all analysis scripts and the BORIS project can be found on the Github page of the main author (https://github.com/porseleyn/UF_Thesis_Analyses).

Table 3: *Annotation definitions.*

Definition descriptions
1. Under a pass falls a score, a score attempt, a successful pass, a forced turnover, and an unforced turnover.
2. Under a successful pass falls a successful pass, and a score.
3. After a point, offense starts after the pull when the disc lands (or is caught in hands).
4. A bricked pull is picked up at the RZ_Def.
5. A pull that lands in the defense's end zone and gets brought to the front line of the end zone is picked up at the RZ_Def.
6. After a turnover, wait a bit before pressing the buttons to change the team on offense so that the turnover is taken within the state of the right team.
7. If there is a violation and the disc goes back, restart the pass count from there, so remove the last pass(es).
8. If a violation discussion takes a long time, do not annotate is as a timeout. Timeout are only signified by the teams during play.
9. A stall-out violation is counted as a forced turnover.
10. If the disc during a score attempt lands outside of the field, select as field location the horizontal zone where it landed. If it lands further than the end zone, select still RZ_Off.
11. If a score attempt is made and there is a contested violation or a retracted point, then it is explicitly a score attempt (so no forced/unforced turnover).
12. When there is a score, there is no need to also annotate a successful pass.
13. After a score and pressing the buttons to change the team on offense, wait a bit before registering the pull pickup so that the pull is taken within the state of the right team.
14. First half ends after the last score before half time. Second half starts when the teams are getting ready for the next point in the second half.

6.3 A look at our data

Each game contained 445 to 748 datapoints (N = 16, M = 591 datapoints). The average time that each team was on offense is 40.4 minutes. Overall, the first half took 47.5 minutes and the second half took 44.1 minutes. On average, the total gametime was 95.8 minutes for each game, and 91.4 minutes excluding the breaks (half time and timeouts).

On average each team scored 12.50 points (SD = 2.97), and made 17.59 scoring attempts (SD = 3.17). For each indicator, the mean and standard deviation was calculated; score efficiency (M = 71.45; SD = 14.85), disc possession (M = 44.27; SD = 7.95), passing accuracy (M = 91.91; SD = 3.69), forced turnovers (M = 6.91; SD = 2.52), unforced turnovers (M = 8.97; SD = 3.70), total turnovers (M = 15.88; SD = 4.83), subs (M = 15.66; SD = 2.73), possession efficiency (M = 8.38; SD = 4.38), passing rate (M = 5.46; SD = 1.43), possession duration (M = 86.23; SD = 18.39). Eight out of 16 games resulted in an early win. On average there were more unforced than forced turnovers.

To check whether the distribution of the differences between the two paired values for each indicator are normally distributed, histograms were plotted and Shapiro-Wilk tests were done. To check for influential outliers, the IQR (Interquartile Range) method was applied. Six outlier values were identified but not removed, and to assess their impact, paired sample t-tests were conducted both with and without outliers for comparison. These are discussed at the appropriate point within the remainder of this thesis. All the indicators passed the other assumptions for doing a paired sample t-test.

7 WHAT MAKES A WINNER

In this section we will discuss the results from all the analyses that were performed, i.e., comparing winning and losing teams, looking at points win margins for each indicator, and looking at the state transition probabilities. For the analyses in this thesis, statistical significance is assumed when $p < 0.05$ (two-tailed) and indicated with *. If $p < 0.01$, then this is indicated with ** (two-tailed).

7.1 Comparison between winning and losing teams

A comparison was made between winning (n = 16) and losing teams (n = 16) by performing paired sample t-tests. For each performance indicator, the mean, standard deviation of winning teams and losing teams, and mean differences and its standard deviations are given, together with t-test values, p-values and effect size values (d), as presented in Table 4. We found out that early win is rather an interaction effect than a predictor, thus we did not include early win in the analyses due to its nature. Effect sizes are considered according to Cohen's D benchmarks (Cohen, 1992), with $d > 0.2$ being small, $d > 0.5$ being medium, and $d > 0.8$ being large. Boxplots of all the indicators can be seen in Figure 5. The indicators that were found to be significantly different between winning and losing teams also had a large effect size. These indicators are attempted scores** (p = 0.003; d = 1.09), score efficiency* (p = 0.015; d = 1.03), disc possession** (p = 0.002; d = 1.66), passing accuracy** (p = 0.009; d = 1.02), total turnovers** (p = 0.008; d = 1.04), number of subs* (p = 0.021; d = -0.82), possession efficiency* (p = 0.023; d = -0.50) and passing rate** (p = 0.001; d = -0.83).

The indicator of passing accuracy had one outlier (on the lower end for losing teams), the indicator of subs had two outliers (on the lower end for winning teams) and the indicator of possession efficiency had three outliers (on the higher extreme for losing teams). For the three indicators, the removal of outliers had minimal effect on the overall pattern of results. The consistency of the results after outlier removal suggests the findings are robust and not driven by a few extreme cases.

7.2 Points win margins

As an ad hoc analysis, we wanted to visualise how the performance indicators relate to the margin of victory. For this, we calculated the point difference between the winning and losing team for each game, called the points win margin. Then, we calculated the indicator margin for wins, i.e., the difference of the indicator value for the winning and losing team. This gave us two points (points win margin and indicator margin for wins) which we could then plot against each other for each game, and we repeated this for each indicator, as can be seen in Figure 6.

The correlation values (r) were also calculated. For Pearson's r correlation interpretation, we consider $r > 0.7$ being strong and $r > 0.9$ being very strong. Significant correlations are score efficiency

Table 4: ***Performance indicators for winning teams (N = 16) and losing teams (N = 16), with paired sample t-test results. The difference is always calculated by subtracting the losing team from the winning team.***

Indicator	Winning teams		Losing teams		Differences		Paired Sample t-test		
	M_W	SD_W	M_L	SD_L	M_Δ	SD_Δ	$t_{(15)}$	p	d
1. Attempted scores	19.13	2.50	16.06	3.09	3.06	3.40	3.608	0.0026**	1.09
2. Score efficiency	78.32	11.36	64.57	15.02	13.76	19.94	2.759	0.0146*	1.03
3. Disc possession	49.36	6.54	39.18	5.69	10.18	10.92	3.730	0.0020**	1.66
4. Passing accuracy	93.60	2.31	90.21	4.08	3.38	4.53	2.988	0.0092**	1.02
5. Forced turnovers	5.94	2.67	7.88	2.00	-1.94	3.00	1.169	0.2607	0.46
6. Unforced turnovers	8.06	3.97	9.88	3.28	-1.81	2.86	-0.433	0.6710	-0.18
7. Total turnovers	14.00	4.50	17.75	4.52	-3.75	2.62	3.060	0.0079**	1.04
8. Number of subs	15.88	1.78	15.44	3.48	0.44	3.93	-2.584	0.0208*	-0.82
9. Possession efficiency	9.38	4.04	7.38	4.60	2.00	6.84	-2.538	0.0227*	-0.50
10. Passing rate	5.33	1.41	5.59	1.47	-0.26	2.37	-5.724	<0.0001**	-0.83
11. Possession duration	94.76	16.63	77.69	16.33	17.08	22.32	0.445	0.6626	0.16

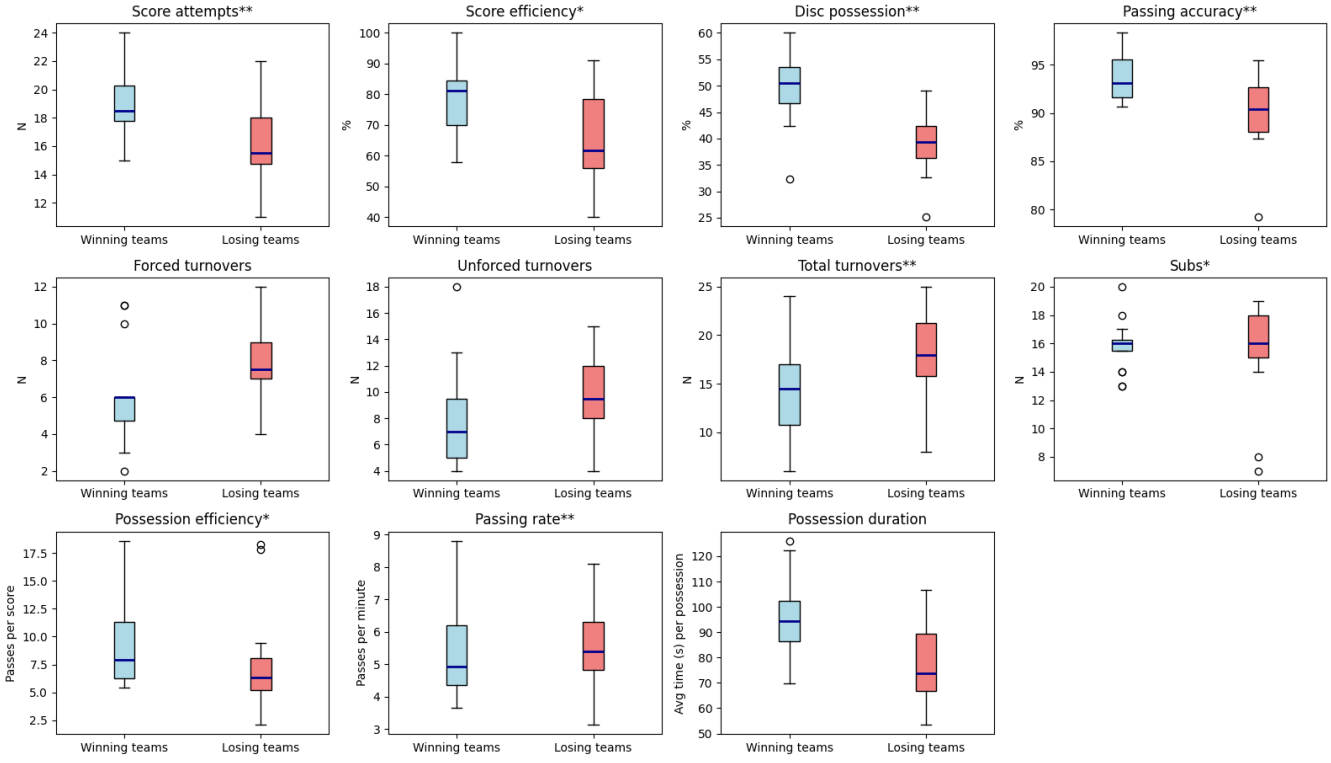

Figure 5: Boxplots of indicators.

and total turnovers, as indicated by asterisks in Figure 6. According to these benchmarks, score efficiency** has a strong positive correlation with points win margin (r = 0.72, p = 0.002), and total

turnovers** has a very strong negative correlation with points win margin (r = -0.99, $p < 0.001$).

7.3 State transition probabilities

A probability matrix shows the probability of a disc being thrown to location Y, given that it was thrown from location X. To calculate state transition probabilities, we divided the passes from field location X to field location Y by the total passes that were thrown from field location X. Figure 7 shows the probabilities of winning teams throwing a disc from location X to location Y, indicated on the x- and y-axis. Moreover, the pull, earned turnover, turnover and score probabilities can also be seen. Figure 8 shows those probabilities for losing teams.

In the probability matrix of the winning teams, we see that most throws were done within the same field location (defensive and offensive middle and red zones). Most points were scored from the offensive red zone (22.04%), followed by the offensive middle zone (3.92%) and defensive middle zone (2.83%). We also see that from the defensive end zone, most throws were thrown to the defensive red zone (48.24%) or end zone (41.71%). Pulls were most often thrown to the defensive red zone (63.48%). After an earned turnover, most throws were thrown to the defensive red zone (41.03%), however, where the turnover was committed is not taken up into the matrix. We do see that where turnovers took place is approximately equally divided over the playing field.

In the probability matrix of the losing teams, we see similar patterns, with most throws being done within the same field location. As with for winnings teams, most points were scored from the offensive red zone (20.77%), followed by the offensive middle zone (5.24%) and defensive middle zone (2.27%). This difference between scores from the offensive and defensive middle zone is however bigger for losing teams than for winning teams. Same as for winning teams, pulls were most often thrown to the defensive red zone (60.81%), but landed approximately equally often in the defensive end (18.92%) and the middle zone (16.67%). After earned turnovers the disc was most often thrown to the defensive red zone (38.89%), followed by the defensive (21.43%) and offensive middle zone (19.05%).

These differences are more clearly shown in the raw count difference matrix in Figure 9. In this matrix, the differences of absolute values (i.e., number of throws at each position) between winners and losers can be seen. What is striking is that winning teams throw far more throws within the offensive red zone than losing teams. Moreover, winning teams threw more passes from the defensive to the offensive middle zone and from the offensive middle to the red zone. However, losing teams have more passes within the defensive red zone than winning teams. Another interesting observation is that, on average, losing teams threw more pulls into the defensive red and end zone than winning teams. This could be due to the fact that after a point, the other team pulls. This means that losing teams had more opportunities to pull, thus explaining the red boxes in the upper row of the matrix.

8 DISCUSSION

This thesis was carried out to answer the question of what performance indicators are the most predictive of success in European ultimate frisbee games. The following section discusses the results, its implications, limitations and recommendations that follow from it.

8.1 Interpretations

The results already give some insight into what indicators are significantly different between winning and losing teams, but it is rewarding to look into what these differences actually mean to the game of ultimate.

8.1.1 Comparison between winning and losing teams. Attempted scores was found to be significantly different** between winning and losing teams, which is in line with the results found in the review study by Lepschy et al. (2018). The more scoring attempts, the more scoring opportunities. If a team has a bid on a score, it means they saw a clear possibility to get past their defense and make the score, even it turns out to be unsuccessful. A higher number of attempted scores for winning teams can be indicative of confidence in throwing execution, as more bids for a score were made in general for winning teams. It can be speculated that winning teams are more assertive on offense than losing teams, and thus have more chances to score.

We found that winning teams had a significantly higher score efficiency*, i.e., the percentage of points scored over total scoring attempts, than losing teams. This is in line with what we found in a lot of other sports and studies (Leicht et al., 2017; Lepschy et al., 2018). Score efficiency is suggestive of being efficient in converting their scoring opportunities, for example by having good throwing connections and flow among team members. This means that players within a team know how to play together, read each other's actions well and act upon them usually caused by a lot of mutual experience, that is, training together.

The difference in disc possession, i.e., the percentage of time spent on offense over time spent on defense, between winning and losing teams was found to be significant**. Just as was found in the study by Lepschy et al. (2018), but then for ball possession in football. This makes sense since the longer a team is on possession, the less their opposition is on offense. Being on offense, there is more chance on scoring, whilst minimising scoring chances for their opponent. Being able to keep control of the disc within a team, and thus control of the game, can be indicative of a strong offensive play, excellence in connecting throws (having throws arriving at the right person or otherwise being saved by your team) and/or an opponent that is weaker on defense than the other team. Longer possession of the disc means more control of the tempo of the game, and can also mentally wear down the opponent.

Winning teams were found to have a significantly higher passing accuracy** than losing teams. Passing accuracy, which is the percentage of completed to toal pass attempts, was often found to be significant in football studies (Lepschy et al., 2018). A higher passing accuracy thus means that winning teams were more efficient at throwing and threw fewer discs away than losing teams. Winning teams are thus better at keeping control of the disc.

Winning teams had fewer total turnovers** committed to them than losing teams. Defensive actions in general are also often indicative of success in basketball studies (Çene, 2018; Leicht et al., 2017; Plakias et al., 2024). The total amount of turnovers gives space

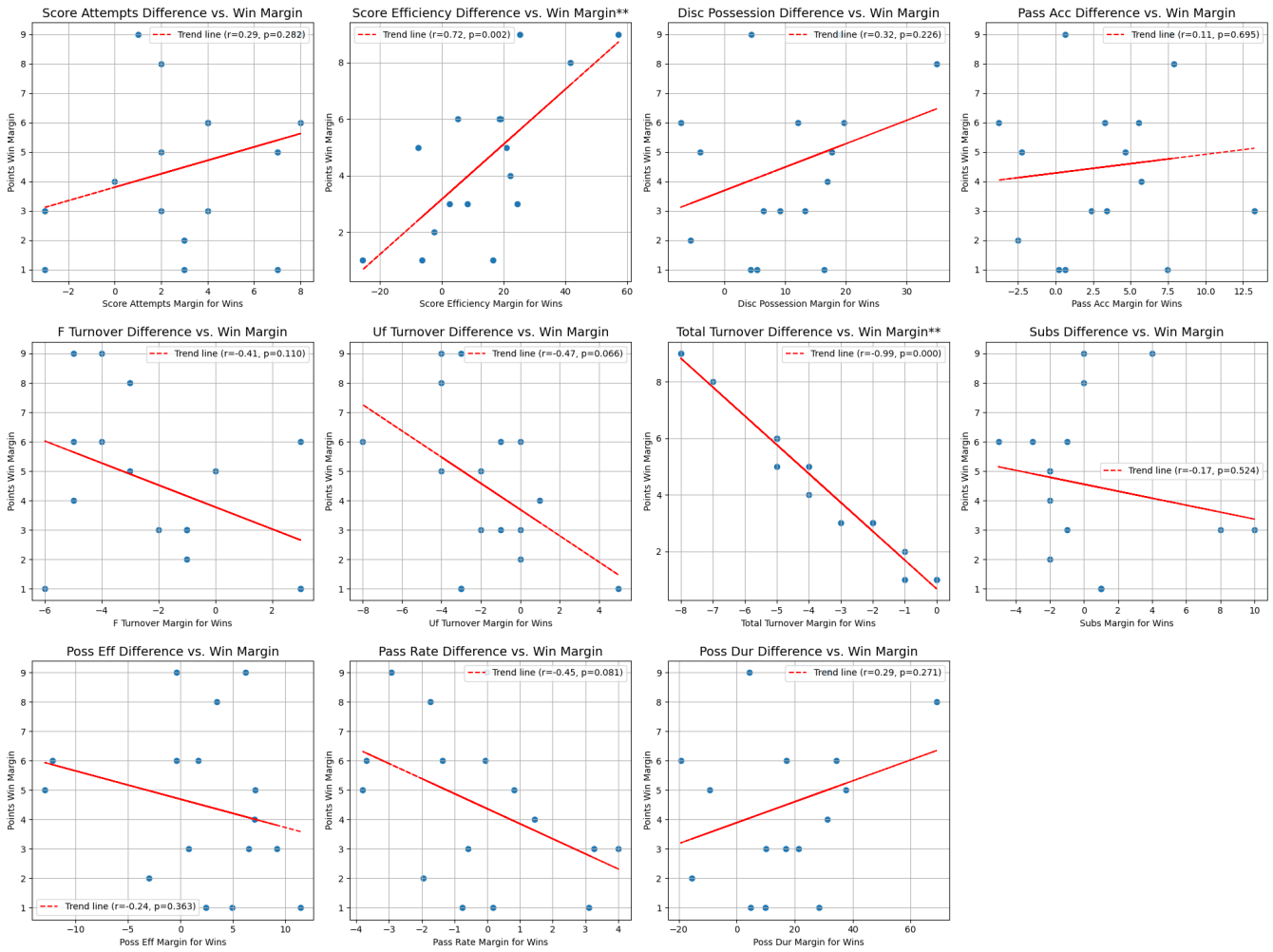

Figure 6: Correlation plots of indicator point difference between winning and losing teams (horizontal axes) and points win margin (vertical axes).

to speculate about certain qualities of the winning teams, which are in line with the qualities that Lam et al. (2021) found in winning teams in their study. The total amount of turnovers that a team has committed can be an indication of strong defense tactics and concentration during a game. Winning teams are speculated to choose better timing, spacing and handler resets, resulting in fewer risky choices. However, this can be dependent of whether the win feels secure, given that being ahead in the game can give more head space to make better decisions than when there is stress about losing the game. In addition, it could be reflective of the winning teams having better conditioning and focus late in games when fatigue can lead to mistakes, thus to winning teams being physically superior to losing teams. It can be an indication of the offensive team maintaining possession more reliably, and in general it means that there were fewer throwaways, drops, or poor decisions, indicating higher thrower skill, awareness, and chemistry with cutters. Turnovers also disrupt the flow and momentum of the offensive team, and thus less turnovers gives way to a steadier flow. What can not be deriven from this research but can be speculated, is that winning teams might have faster or more agile players that are harder to pressure into making mistakes. Everyone makes mistakes, therefore a focus on maximising forced turnovers is suggested to be more valuable than minimising unforced turnovers. More forced turnovers endured by a team and unforced turnovers that are at their own fault signifies more opportunity for the other team to spend time on offense, which is the optimal position to be in to score points.

In terms of substitute players, winning teams were found to have significantly less subs* than losing teams. A lower number of subs indicates smaller rotations, thus the remaining players each having more playing time. This could have benefits, namely that a team knows each teammate's playing style better, allowing for a better chemistry and connections. A cohesive lineup translates to consistency and familiarity on the field. Moreover, a player's body stays warm if they play a lot, allowing for more intense activity. On top of that, playing more could also help with being more in

Figure 7: Passing behaviour probability matrix of winning teams.

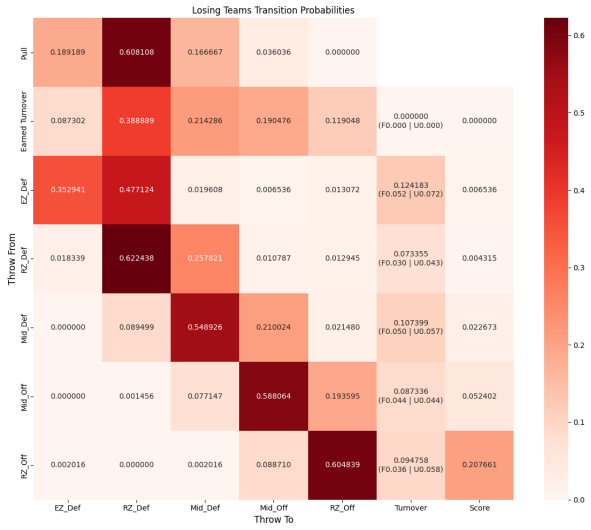

Figure 8: Passing behaviour probability matrix of losing teams.

the game mentally, possibly leading to a better understanding of the opponent's play style and adjusting to that.

For possession efficiency, i.e., the number of passes in each successful scoring attempt, winning teams had a significantly higher value* than losing teams. This means that winning teams make more passes when they score, which suggests that winning teams are more patient and deliberate with their disc possession, working the disc carefully down the field rather than rushing to score. This said, losing teams might be trying to score quickly rather than build structured possessions, with fewer passes per score possibly indicating more aggressive or hurried attacks. Having fewer resets

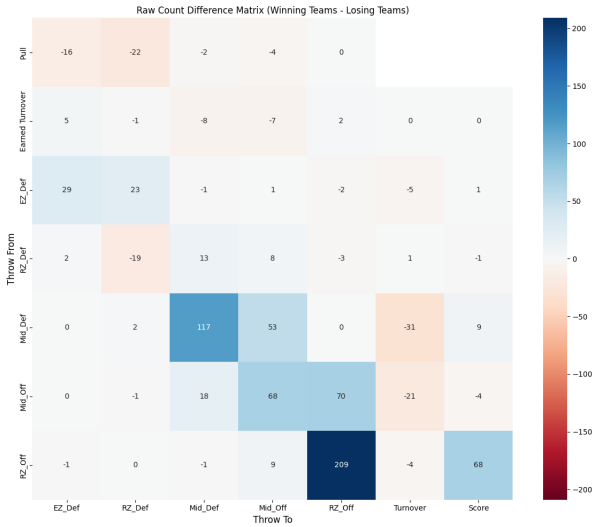

Figure 9: Passing behaviour raw count matrix of differences (winning teams - losing teams).

could reflect a lack of patience or decision-making under pressure, possibly relying on individual plays or lucky breaks.

Passing rate, i.e., the number of passes made per minute during possession, was significantly lower** for winning teams than for losing teams. In other words, winning teams seem to be more patient and thoughtful with the disc than losing teams. Advancing the disc with fewer passes could point to several reasons, such as needing less handler resets and having more efficient cuts and throws. On the other hand, losing teams making more passes per minute can be indicative of struggles to make good cuts, causing a higher passing rate to keep resetting the disc and stall count. Or, losing teams playing more rapidly could be due to pressure leading to riskier decisions and more errors. However, a lower pass rate does not mean less possession, but rather a more productive possession: winning teams pass less frequently over time, taking their time to make the right pass, not just any pass. This sounds contradictory to possession efficiency, but passing rate is about pace, while possession efficiency is about quality.

Taken all together, it can be derived from the results that winning teams consistently demonstrate assertiveness on offense, high team chemistry, control-oriented, risk-avoident and patient play, strong defense and good spatial awareness. This can be seen in a higher number of attempted scores, lower turnover and passing rates, greater overall score efficiency, disc possession, passing accuracy, possession efficiency, and a lower number of subs. Together, these indicators point to a style of play that emphasises disc control, patience, strategic offense, and defensive pressure that limits opponent opportunities, of which controlled disc possession and strong defense are in line with the results found by Lam et al. (2021).

Interestingly, in some sports such as football and rugby, offensive tactics seemed to be associated more with winning (Lepschy et al., 2018; Woods et al., 2017), while in other sports such as basketball, a mix of both offensive and defensive tactics were more indicative

of success (Çene, 2018; Leicht et al., 2017; Plakias et al., 2024). In ultimate, it also seems to be a mix of both.

8.1.2 Points win margins. We see that score efficiency** has a large positive effect size, which means that the larger the point difference, the larger the difference in score efficiency between winning and losing teams. As explained in section 8.1.1, score efficiency is suggestive of being efficient in converting scoring opportunities, and indicates good throwing connections and flow among team members. The more efficient a team is at scoring when they have the chance, the more likely they are not just to win — but to dominate.

Total turnover** has a large negative effect size, meaning that the larger the point difference, the fewer turnovers were committed on the winning team, i.e., the losing team had endured more turnovers than winning team. In other words, teams that scored more were also better at committing turnovers. Fewer turnovers signifies better possession of the disc, which in turn paves the way for more chances for scoring. In terms of turnovers it can signify that winning teams were better at both holding possession through active offense, making fewer mistakes and gaining possession through active defense.

Taken together, score efficiency can be suggestive of a good throwing connections and flow among team members, while total turnovers indicate that the greater the point difference, the better the winning team was in holding possession of the disc and thus having scoring opportunities and in committing to an active defense to gain possession of the disc. How efficient a team is on offense and how often they give up the disc strongly relate to how much they win or lose by - it helps to explain the scale of victory or defeat.

8.1.3 State transition probabilities. Next, we interpret what the information that can be read from the matrices really means. The winning and losing matrices tell you the probability of given the disc was thrown from location X, that it went to location Y.

It can be clearly seen in the winning and losing teams' matrices (Figures 7 and 8) that the offensive red zone was the most successful location to throw from for scoring a point. This is the zone that is closest to the offensive red zone, so it makes sense that this is the most thrown from location given that it requires the least distance to be covered.

What is interesting when comparing the probability matrix of the winning teams with that of the losing teams, is that losing teams endure more turnovers within their defensive end zone. Losing possession of the disc in front a defensive end zone gives a great opportunity for the opposition, who then do not have to cross the whole field to score a point. As established by the indicators, losing teams also have more turnovers in general than winning teams, but what is interesting to see is that the difference between losing and winning team turnovers is biggest in the defensive middle zone. This could be because that is where the losing teams loosen their control over the disc more than in the other zones, since it might feel like a "safe" zone, compared to the risky defensive red zone, or the offensive part of the field where action needs to be taken to score a point. Lam et al. (2021) found that winning teams caused more turnovers in the middle zone than losing teams, and with the two middle zones taken together, we find the same result.

Losing teams threw more pulls deep into the field than winning teams. However, this could be due to losing teams having to perform more pulls, since after a point is scored, the losing team has to pull. Pulling behaviour is trainable and does not need to say much about a game, except a deeper and floatier pull gives the defense enough time to set up.

Winning teams made more passes within the zones, and to the next or previous zones than losing teams. They make thus more shorter passes than losing teams, as apparent by the blue squares of the differences matrix, where most squares around each zone are light shades of blue. In the study by Lam et al. (2021) they found the same result. Moreover, losing teams throw more often to field locations further away from the origin, as can be seen by the negative numbers in the left bottom and upper right squares in the differences matrix. In other words, while winning teams seem to keep the disc closer to its current field position, losing teams make longer throws, spanning over multiple zones. This could mean losing teams often used a tactic that was less successful, or had a hard time breaking the defense of the opposition and therefore had to throw a lot more deep throws, which are often with less accuracy and more risky.

Losing teams have more passes within their defensive red zone than winning teams. This could indicate that they had a hard time getting out of their defensive red zone. On offense, playing near your end zone can be stressful, since making a drop gives the oppostion a huge advantage to scoring a point. It is therefore better to play safe than risking throwing it deep from one's own red zone as to not make mistakes, which could be the case for the losing teams.

The opposite can be seen for winning teams. They made the most passes within the offensive red zone, suggesting that they made more passes as to keep control of the disc, waiting for an opportunity to break the defense and score a point.

All with all, the state transition probabilities show that winning and losing teams have similar strategies, such as scoring most from the offensive red zone, yet they apply different tactics during the game. Losing teams endured most turnovers in their defensive endzone, which is a huge disadvantage. Even though losing teams threw more pulls deep into the field than winning teams but it did not make them win. The difference between losing and winning team turnovers is the biggest in the defensive middle zone, likely due to the losing teams loosening the reigns on the disc whilst being in a more comfortable zone than the other zones. It can be speculated that losing teams make longer throws than winning teams, while winning teams keep the disc within and around a zone and advance slowly but steady. On one hand, losing teams have more passes within their defensive red zone than winning teams, suggesting they had a hard time getting out of their defensive red zone. On the other hand, winning teams made the most passes within the offensive red zone, suggesting that they made more passes as to keep control of the disc, waiting for an opportunity to break the defense and score a point. This means that most of the action happened in the red zones, as opposed to most action happening in the middle zone in the study by Lam et al. (2021).

8.2 Implications

This study provides guidelines for successful strategies and tactics in ultimate frisbee games, which can be picked up by teams or worked through to increase offensive and defensive plays within

teams. The results are up for interpretation, but through experience with the sport we have interpreted them as best as possible. The state transition probability analysis part of this research might impact the field of sports analysis, by serving as an example of how field locations in ultimate frisbee or other sports can be used to interpret team plays.

Potential consequences that may arise from this study's findings are that these results are generalised over all of ultimate, while it focuses solely on games used in European contexts. Other continents can use different game plays, so it would be interesting to see if the same results can be seen in say, American or Asian game contexts.

Ultimate frisbee is an under-researched sport. By focusing on European ultimate, we contribute to filling this research gap and providing a more comprehensive understanding of the sport globally. Moreover, this study differs from others mainly because performance analysis has only been done once before in ultimate frisbee on a different population. This analysis significantly contributes to the sports analytics of ultimate frisbee, as it is only the second attempt at studying performance indicators in this sport. Moreover, the findings of this research reveal how performance indicators differ between the American and European leagues. These observations enhance our understanding of how the competitive environment can display different strategic and performance challenges in ultimate frisbee.

8.3 Limitations

Considering the scarcity of studies that address performance indicators in ultimate frisbee, the main objective of this study was to identify the strongest predictive variables of winning European ultimate frisbee games at the elite level. Due to the scarcity of research in ultimate frisbee, there is not a generally utilized methods, so in this study, methods are used that are seen in related sports such as football, basketball and rugby. This could have led to certain sport-specific nuances that went lost in the conversion to ultimate specific indicators, which may be the reason why some indicators did not prove predictive.

The recorded games that were used for annotation were edited, not raw footage. Edited as in, they were used for broadcasting, which sometimes included replays after important events, creating blanks in the overall overview of the game. In the third game (Ireland against Italy), footage of the first point was missing, and sometimes after a half time, a timeout or a replay of a critical moment, the first few seconds of footage were missing. Moreover, as mentioned before, in some games the lines separating the middle and red zones were unclear. This led to some suggestive interpretation of the lines by the annotater. Moreover, in one game it rained, leading to low quality footage due to raindrops being on the camera.

Sometimes it was hard to distinguish forced from unforced turnovers, for example where a disc was being tapped away by the defense, but it was unreachable for the offensive player anyways, or where the players obstructed the view on the disc. These ambiguous situations were noted down and added to the annotation definitions to ensure reliable annotating where possible. Moreover, the expertise of the annotater enabled ambiguous situations to be properly handled and annotated accordingly.

What was not taken into account were the weather conditions. Rain can make the disc and field surface slippery, while wind allows for different strategies to be employed to counteract the natural forces. Extreme temperatures can tire out players faster. However, since both teams play in the same conditions, it should not give one team an advantage over the other team. The teams would both have to adapt their strategies. Another interacting factor that was not taken into account is the home or away advantage. Experiencing the same limitation as referred to by Lam et al. (2021), the sample pool was quite small, with only three out of 16 games played by the "home" team. However, the field of playing is not necessarily the home field of the "home" team, only the country is. Since spirit is quite a unique part of ultimate, it would also be interesting to see if that could be applied as an (interaction) variable.

A possible limitation is that the reliability of the novel indicators has not been tested, nevertheless this research can be used as a starting point for acquiring indicators for this specific sport.

8.4 Recommendations

The fact that the indicators of forced and unforced turnovers and possession duration were not found to be significantly different between winning and losing teams could be due to the small sample size, annotation errors or them not being predictive of success. For future research and studies building on this paper, we recommend to increase the sample size. Annotation errors can be checked by doing a reliability check, e.g., redoing the first game annotation to see if it still matches.

We feel that some indicators would predict differently when looking separately at the first half and second half of the game, such as passing rate. We suspect that a change in play style can be activated during the game, so looking at the indicators during different parts of the game could be an interesting take. Moreover, early wins was included as a performance indicator in the analyses but not considered a true performance indicator in ultimate, as the number of early wins for both winning and losing teams is the same and thus cannot be compared. As a result, the early win indicator was not included in the analyses, but it could be interesting to look at early win as an interaction effect between winning and losing teams, for example, comparing winning teams with early win against winnings teams without an early win, and the same for losing teams. Another interesting take would be consider close (i.e., small score differences) and unbalanced matches (i.e., large score difference), as was done in the studies by Çene (2018) and Lepschy et al. (2018). They suspect that in close matches, both teams work hard to win, while in unbalanced matches the game could already be decided and different strategies may be at play.

9 CONCLUSION

This study aimed at examining performance indicators that are most predictive of success in European ultimate frisbee games. For this, a set of metrics was designed, which were then extracted from existing recorded ultimate frisbee games from the EUC 2023, followed by analyses that revealed what performance indicators have good predictive power for success. Moreover, the field locations were kept into account to gain more insight into what passing behaviour was used by winning and losing teams.

In conclusion, the indicators that proved to be significantly different between winning and losing teams were attempted scores, score efficiency, disc possession, passing accuracy, total turnovers, number of substitutions, possession efficiency and passing rate. These indicators point to a style of play that emphasizes disc control, patience, strategic offense, and defensive pressure that limits opponent opportunities. When looking at points win margin, score efficiency and total turnovers proved to be significant. These indicators suggest that winning teams have better throwing connections and flow among team members and active defense than losing teams. How efficient a team is on offense and how often they give up the disc strongly relate to how much they win or lose by. The state transition probabilities revealed that winning and losing teams use similar strategies, such as scoring most from the offensive red zone, yet apply different tactics during the game. Moreover, losing teams make longer throws throws than winning teams, while winning teams keep the disc within and around a zone and advance slowly but steady. Winning teams were characterised by having a patient rhythm in the offensive red zone, while losing teams have been seen to use this strategy in their defensive red zone.

Now, looking back, some things could have been done differently. For future research it is therefore recommended to look into variables that can have interaction effects with the other indicators. Nevertheless, this study adds to the scientific knowledge gap of underresearched sports and has translated gut suspicions to tangible results concerning ultimate frisbee gameplay.

10 ACKNOWLEDGEMENTS

I would like to thank Maarten H. Lamers, the primary supervisor, for giving advice and guidance during the whole thesis process. The energising brainstorm sessions and critical questions helped this thesis develop to its full potential. I would like to thank Tiago G. Russomanno, the secondary supervisor, for his expertise in sports analysis, his experience from working on previous relevant papers and advice on how to approach the state transition probabilities analysis. I would like to thank kreatif.minds, Sam Kehl, and the World Flying Disc Federation (WFDF) for permission to use their photographs in this thesis.

REFERENCES

[1] Jacob Cohen. 1992. A power primer. *Psychological Bulletin* 112 (1992), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

[2] Olivier Friard and Marco Gamba. 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* 7 (11 2016), 1325–1330. https://doi.org/10.1111/2041-210X.12584

[3] Jean Francis Gréhaigne, Paul Godbout, and Daniel Bouthier. 1999. The foundations of tactics and strategy in team sports. *Journal of Teaching in Physical Education* 18 (1999), 159–174. Issue 2. https://doi.org/10.1123/jtpe.18.2.159

[4] Mike D. Hughes and Roger M. Bartlett. 2002. The use of performance indicators in performance analysis. *Journal of Sports Sciences* 20 (10 2002), 739–754. https://doi.org/10.1080/02640102320675602

[5] Carlos Lago-Peñas, Miguel Gómez-Ruano, Diego Megías-Navarro, and Richard Pollard. 2016. Home advantage in football: Examining the effect of scoring first on match outcome in the five major European leagues. *International Journal of Performance Analysis in Sport* 16 (2016), 411–421. https://doi.org/10.1080/24748668.2016.11868897

[6] Hilary Lam, Otto Kolbinger, Martin Lames, and Tiago Guedes Russomanno. 2021. State Transition Modeling in Ultimate Frisbee: Adaptation of a Promising Method for Performance Analysis in Invasion Sports. *Frontiers in Psychology* 12 (5 2021), 664511. https://doi.org/10.3389/fpsyg.2021.664511

[7] Anthony S. Leicht, Miguel A. Gomez, and Carl T. Woods. 2017. Team performance indicators explain outcome during women's basketball matches at the olympic games. *Sports* 5 (12 2017). https://doi.org/10.3390/sports5040096

[8] Hannes Lepschy, Hagen Wäsche, and Alexander Woll. 2018. How to be Successful in Football: A Systematic Review. *The Open Sports Sciences Journal* 11 (7 2018), 3–23. Issue 1. https://doi.org/10.2174/1875399x01811010003

[9] Felicity Lord, David B. Pyne, Marijke Welvaert, and Jocelyn K. Mara. 2020. Methods of performance analysis in team invasion sports: A systematic review. *Journal of Sports Sciences* 38 (10 2020), 2338–2349. https://doi.org/10.1080/02640414.2020.1785185

[10] Kazimierz Mikolajec, Adam Maszczyk, and Tomasz Zajac. 2013. Game indicators determining sports performance in the NBA. *Journal of Human Kinetics* 37 (2013), 145–151. https://doi.org/10.2478/hukin-2013-0035

[11] Elia Morgulev and Felix Lebed. 2024. Beyond key performance indicators: Theoretical-methodological discussion of performance analysis (sports analytics) research. *German Journal of Exercise and Sport Research* 54 (2 2024), 1–6. https://link.springer.com/article/10.1007/s12662-024-00944-8

[12] Peter O'Donoghue, Anna Mayes, Kate M. Edwards, and Jess Garland. 2008. Performance Norms for British National Super League Netball. *International Journal of Sports Science Coaching* 3 (12 2008), 501–511. https://doi.org/10.1260/174795408787186486

[13] Spyridon Plakias, Christos Kokkotis, Dimitrios Pantazis, and Themistoklis Tsatalas. 2024. Comparative analysis of key performance indicators in Euroleague and national basketball leagues. *Journal of Physical Education and Sport ® (JPES)* 24 (2024), 1360–1372. Issue 6. https://doi.org/10.7752/jpes.2024.06154

[14] Carl T. Woods, Wade Sinclair, and Sam Robertson. 2017. Explaining match outcome and ladder position in the National Rugby League using team performance indicators. *Journal of Science and Medicine in Sport* 20 (12 2017), 1107–1111. https://doi.org/10.1016/j.jsams.2017.04.005

[15] World Flying Disc Federation. 2025. WFDF Publishes 2025–2028 Ultimate Rules. https://wfdf.sport/2025/01/wfdf-publishes-2025-2028-ultimate-rules/

[16] Francisco Manuel Ávila Moreno, Luis Javier Chirosa-Ríos, Aurelio Ureña-Espá, Demetrio Lozano-Jarque, and David Ulloa-Díaz. 2018. Evaluation of tactical performance in invasion team sports: a systematic review. *International Journal of Performance Analysis in Sport* 18 (3 2018), 195–216. https://doi.org/10.1080/24748668.2018.1460054 Review: handig!

[17] Erhan Çene. 2018. What is the difference between a winning and a losing team: insights from Euroleague basketball. *International Journal of Performance Analysis in Sport* 18 (1 2018), 55–68. https://doi.org/10.1080/24748668.2018.1446234

Appendix A GAMES USED IN ANALYSIS.

Table A: ***Games overview with teams, results, dates, and video links.***

Game	*Team A*	*Team B*	*Result (A-B)*	*Date*	*URL*
1	Poland	Lithuania	15–14	16-07-2023	link
2 (Final)	France	Italy	15–7	22-07-2023	link
3 (Semi-final)	Ireland	Italy	14–15	21-07-2023	link
4 (Quarter final)	Latvia	Ireland	13–15	20-07-2023	link
5	Austria	Finland	11–15	19-07-2023	link
6	Great Britain	Germany	15–12	19-07-2023	link
7	France	Lithuania	15–6	19-07-2023	link
8	Denmark	Switzerland	11–10	18-07-2023	link
9	Austria	Slovakia	15–9	18-07-2023	link
10	Czech Republic	Germany	6–15	17-07-2023	link
11 (Bronze medal)	Ireland	Great Britain	15–10	21-07-2023	link
12 (Semi-final)	France	Great Britain	15–10	21-07-2023	link
13 (Pre-quarter)	Poland	Hungary	15–12	20-07-2023	link
14 (Quarter-final)	Germany	France	9-15	20-07-2023	link
15 (Pre-quarter)	Latvia	Finland	15–9	20-07-2023	link
16	Hungary	Italy	12-15	17-07-2023	link

Appendix B ANNOTATION NOTES.

A scanned in image of the annotation notes during preliminary game (Figure A) and consequent games (Figures B and C).

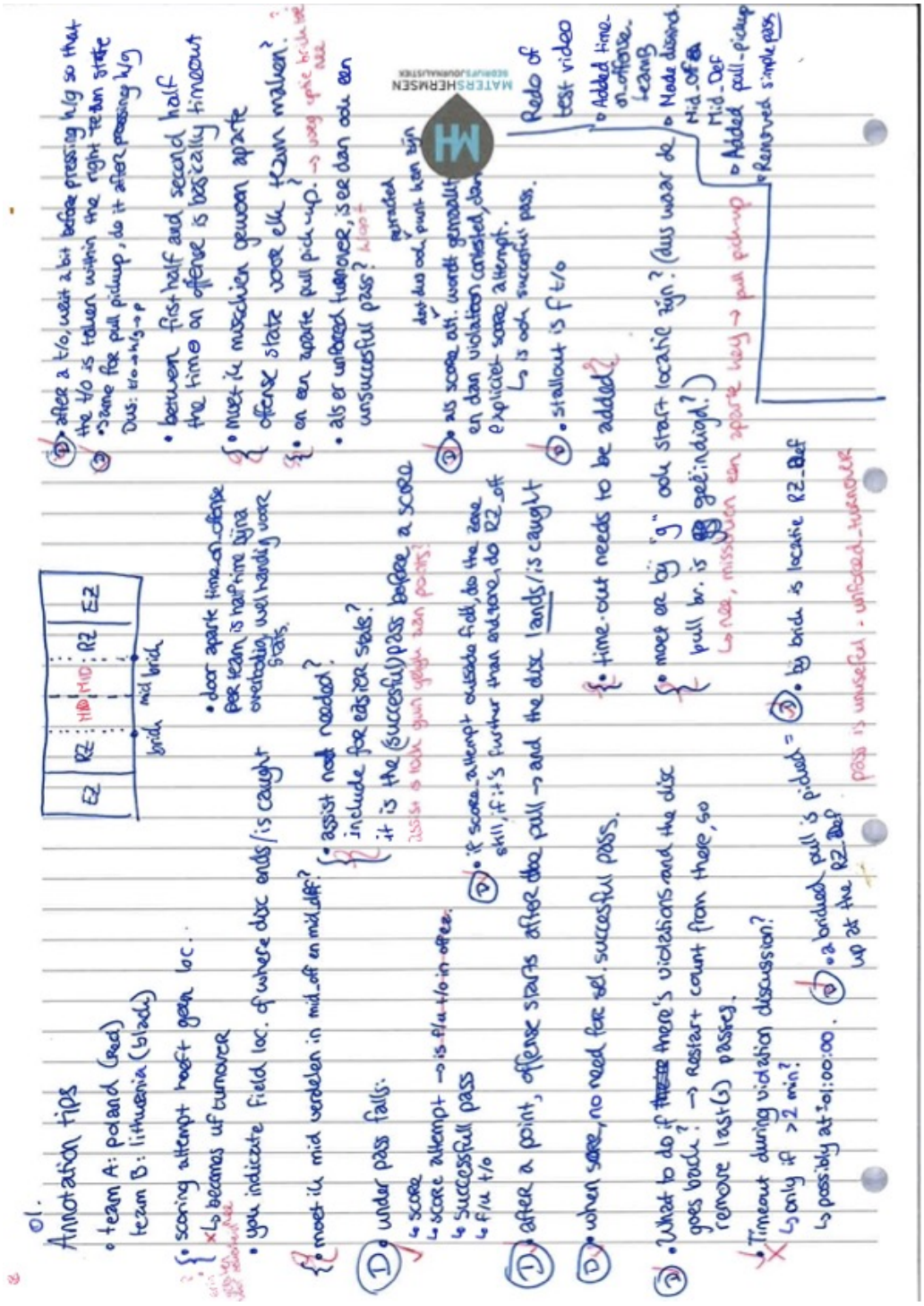

Figure A: General annotation notes.

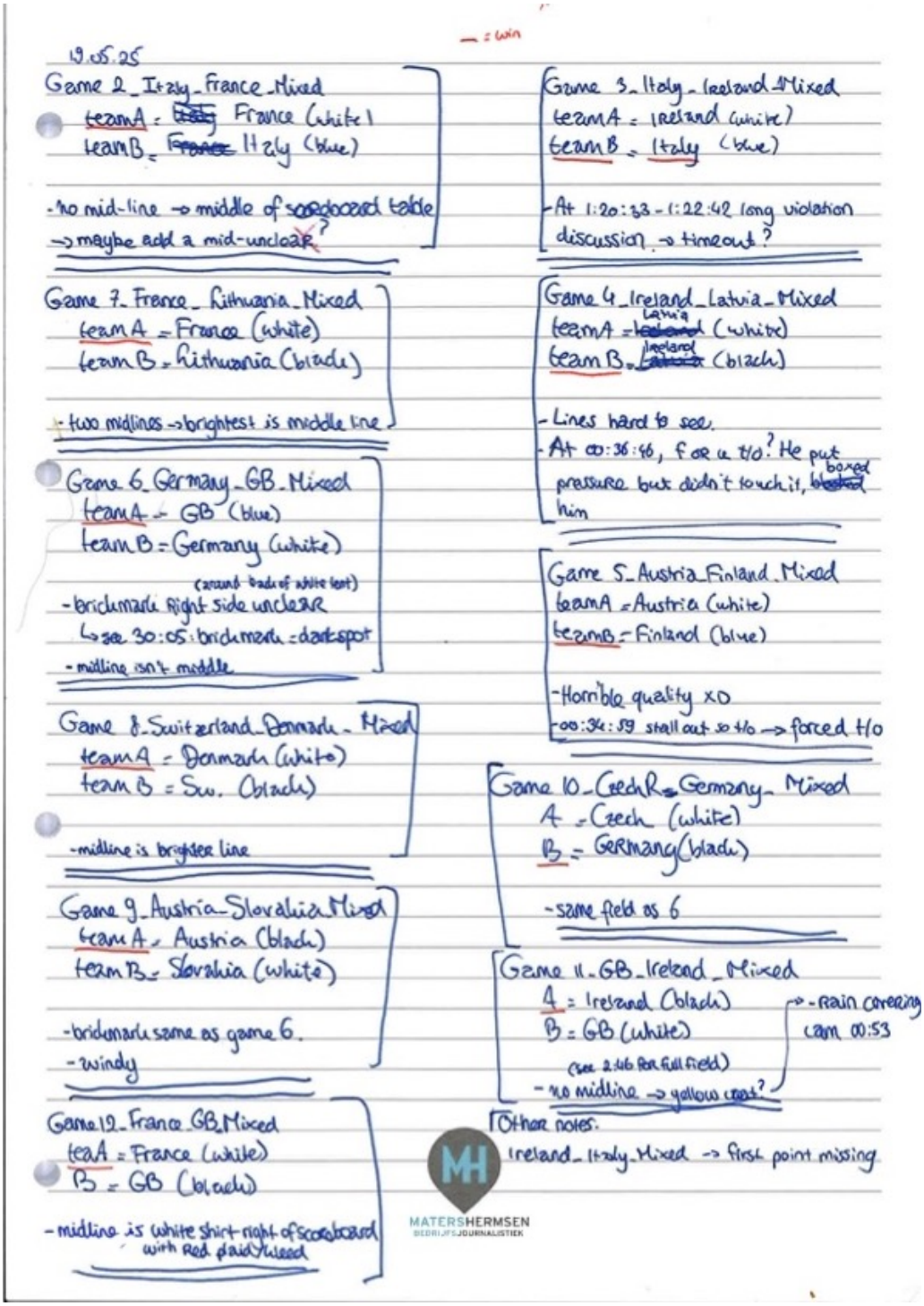

Figure B: Annotation notes per game.

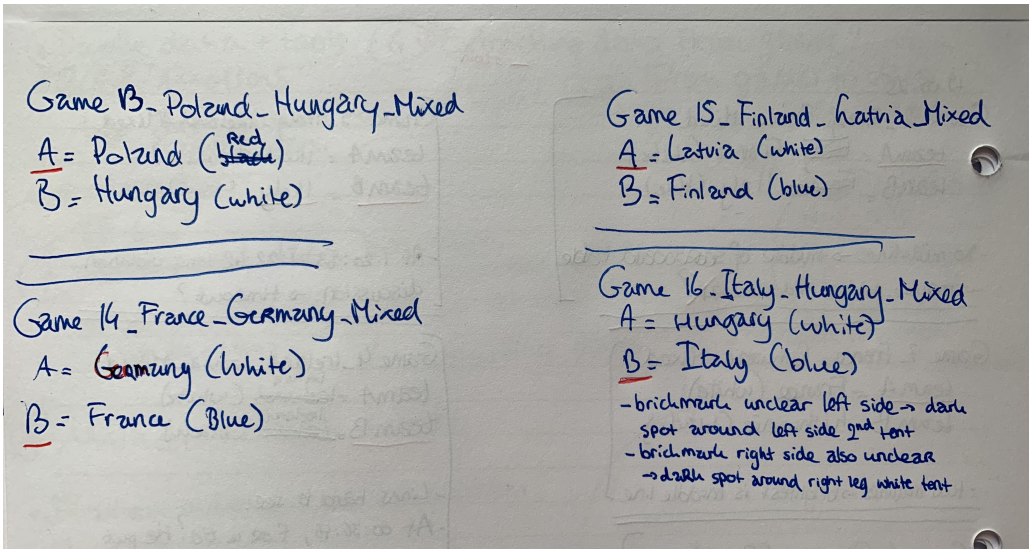

Figure C: Annotation notes per game.