

## Article

# The “Hockey” Assist Makes the Difference—Validation of a Defensive Disruptiveness Model to Evaluate Passing Sequences in Elite Soccer

Leander Forcher <sup>1,\*</sup> , Matthias Kempe <sup>2</sup>, Stefan Altmann <sup>1,3</sup> , Leon Forcher <sup>1,4</sup> and Alexander Woll <sup>1</sup>

<sup>1</sup> Institute of Sport and Sport Science (IfSS), Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany; Stefan.altmann@kit.edu (S.A.); Leon.forcher@kit.edu (L.F.); Alexander.woll@kit.edu (A.W.)

<sup>2</sup> Center for Human Movement Sciences, University of Groningen, University Medical Center Groningen (UMCG), 9713 GZ Groningen, The Netherlands; M.kempe@umcg.nl

<sup>3</sup> TSG ResearchLab gGmbH, 74939 Zuzenhausen, Germany

<sup>4</sup> TSG 1899 Hoffenheim, 74939 Zuzenhausen, Germany

\* Correspondence: Leander.forcher@kit.edu

**Abstract:** With the growing availability of position data in sports, spatiotemporal analysis in soccer is a topic of rising interest. The aim of this study is to validate a performance indicator, namely D-Def, measuring passing effectiveness. D-Def calculates the change of the teams’ centroid, centroids of formation lines (e.g., defensive line), teams’ surface area, and teams’ spread in the following three seconds after a pass and therefore results in a measure of disruption of the opponents’ defense following a pass. While this measure was introduced earlier, in this study we aim to prove the usefulness to evaluate attacking sequences. In this study, 258 games of Dutch Eredivisie season 2018/19 were included, resulting in 13,094 attacks. D-Def, pass length, pass velocity, and pass angle of the last four passes of each attack were calculated and compared between successful and unsuccessful attacks. D-Def showed higher values for passes of successful compared to unsuccessful attacks ( $0.001 < p \leq 0.029$ ,  $0.06 \leq d \leq 0.23$ ). This difference showed the highest effects sizes in the penultimate pass ( $d = 0.23$ ) and the maximal D-Def value of an attack ( $d = 0.23$ ). Passing length ( $0.001 < p \leq 0.236$ ,  $0.08 \leq d \leq 0.17$ ) and passing velocity ( $0.001 < p \leq 0.690$ ,  $-0.09 \leq d \leq 0.12$ ) showed inconsistent results in discriminating between successful and unsuccessful attacks. The results indicate that D-Def is a useful indicator for the measurement of pass effectiveness in attacking sequences, highlighting that successful attacks are connected to disruptive passing. Within successful attacks, at least one high disruptive action (pass with D-Def > 28) needs to be present. In addition, the penultimate pass (“hockey assist”) of an attack seems crucial in characterizing successful attacks.

**Keywords:** soccer; performance indicators; passing; tracking data; spatiotemporal data; sports analytics



**Citation:** Forcher, L.; Kempe, M.; Altmann, S.; Forcher, L.; Woll, A. The “Hockey” Assist Makes the Difference—Validation of a Defensive Disruptiveness Model to Evaluate Passing Sequences in Elite Soccer. *Entropy* **2021**, *23*, 1607. <https://doi.org/10.3390/e23121607>

Academic Editor: Filipe Manuel Clemente

Received: 5 November 2021

Accepted: 26 November 2021

Published: 30 November 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the ongoing digitalization in sports, the importance and usage of data in soccer has grown constantly in the last decades [1,2]. In line with this development, the use of spatiotemporal tracking data for tactical game analysis is seen as a beneficial tool to produce more objective, time-efficient, and in-depth analyses [3,4]. In the past, match analysis focused on the physical and technical performance of soccer players [5,6]. With this development of new analysis methods using tracking data, the tactical performance is analyzed more frequently, for example using measures of collective organization (e.g., centroids, spread measures) [7]. By the analysis of tracking data, the exact positioning of players, their spatial formation, and the inter-player distance can be measured more easily and objective [8]. With it, the dynamic, interactive, and complex nature of soccer can be con-

sidered to better characterize performance in soccer, which was unattended in traditional (e.g., notational) analysis approaches [2].

In this context, the investigation of passing behavior received broad attention as it is the most frequently occurring tactical pattern in a game [9,10] and therefore is one of the key skills in soccer [11]. In addition, effective passes are seen as crucial for successful performance [12] and team success in soccer [13].

In the past, the analysis of passing behavior usually did not go beyond numbers of passes and various forms of completion rates [11] revealing little information about passing performance and overlooking factors such as value or reward of passes. However, with the stated availability of spatiotemporal tracking data more complex investigations and key performance metrics on passing behavior in soccer have been developed in recent years [9,10,12,14].

On one hand, there have been several investigations from a computational view that resulted in various prediction models or the estimation of the risk and reward of a pass [9,10]. On the other hand, there have been investigations from sport science research groups to measure passing effectiveness resulting in key performance metrics such as the effect of passes on outplayed opponents and space control in the final third [12] or *Dangerosity* (a metric quantifying the danger of any attacking action or a whole attack to end in a goal) [14].

While these approaches and performance indicators provide new insights and translate practical ideas into numbers, they mostly also overvalue forward passes as they solely connect effective passing to scoring opportunities and advancing in close proximity of the goal [15]. In doing so, these approaches overvalue specific passes (e.g., assists) that might be crucial for the actual development of a successful attack. To account for the complex dynamics of build-ups and space creation in all areas of the pitch, not only the spaces in front of the goal, Goes et al. [15] introduced a passing evaluation tool called D-Def. D-Def is based on the idea that an effective pass disrupts the opposing defense and creates space to achieve scoring opportunities. In this context, other studies showed a higher chance of scoring against an unbalanced defense compared to a balanced defense [16,17]. Earlier studies already revealed that D-Def can differentiate between good and bad passers on the individual level [15] and is a better predictor of the final result of a match compared to other performance indicators [18]. However, since the end result of a game is strongly influenced by chance [19], the relationship of D-Def and success should not solely be traced back to the end result. Consequently, in this study, we investigate the connection of D-Def with the success of a team in the context of individual attacks.

The evaluation of possessions and attacking sequences and therefore also the evaluation of key performance indicators in soccer is generally based on summation scores or singular events. Following this approach, the majority of published spatiotemporal analysis studies using tracking data reduce their analysis on one single summarized value for a whole game or a possession (e.g., total number of shots on goal, percentage of passing accuracy) (Lepschy, 2021; Dufour, 2017). Other studies compare individual possessions with each other using summarized values [20] or focus on the valuation of single actions like shots on goal (e.g., expected goals) [21,22] or passes [15]. In doing so, they do not reveal evidence about the consecutive chain of actions in the attacking process. The approach of the analysis of consecutive actions is strengthened by Sarmiento et al. [2] who pointed out that the analysis of sequential aspects of the game is important to increase the practical impact of match analysis in soccer. Using this approach, it cannot be discovered which pass in an attack was actually important or if some actions benefit from each other (e.g., dribblings are possibly more effective after switching sides with a long pass across the pitch). Given the complex and partially chaotic structure of soccer, it is important to get insights into the chain of events in an attack (e.g., which action of an attack is most important for success?) to further validate performance indicators, as the actual significance of them is still pending [23].

While the idea of analyzing chains of actions is already established in other sports, like tennis or American football [24,25] (e.g., using Markov chains) it is seldomly used in soccer.

One study using this approach was conducted by Kempe and Memmert [26] investigating consecutive actions in the sequence of an attack to evaluate the influence of creative action on goal scoring. In this study, the authors rated the creativity of the last eight actions of attacks that led to goals in the Football FIFA World Cup 2010 and 2014, as well as the Football UEFA Euro 2016. While this study provided empirical evidence for the importance of creativity in soccer, the authors used a rather simple study design with notational data and only focused on goal scoring attacks.

In the present study, the goal is to establish a new approach for the analysis of attacks by investigating individual consecutive actions (passes) in the chain of an attack using the opportunities of tracking data.

Therefore, the purposes of this study are to use this approach (investigation of consecutive actions in the chain of an attack) (i) to test whether the quantitative pass model, D-Def, is a valid measure for pass effectiveness on a single possession level and (ii) to reveal practice-relevant information about the characteristics of successful passing in the attacking process in elite soccer. To do so, we will differentiate between passes of successful and unsuccessful attacks using a spatiotemporal tool that objectively measures the danger of an attack. Accounting for the first aim, we hypothesized that passes of successful attacks show higher D-Def values compared to unsuccessful attacks. For the second aim, an explanatory approach of the consecutive actions of an attack will be applied.

## 2. Materials and Methods

### 2.1. Data

We used an observational design in which position tracking data and event data of 258 games of the Dutch Eredivisie season 2018/19 were included. Of all 305 games of the whole season, 47 matches (15.4%) were excluded because of erroneous or missing data.

The position data of all 22 players and the ball as well as the associated ball event data was collected by a semi-automatic optical tracking system (TRACAB, Chyron Hego, Melville, NY, USA) that measures the X- and Y-coordinates of all players and the ball with a sampling frequency of 25 Hz. This camera-based tracking system has recently been validated [27]. Before the data processing and analysis, the raw position data were pre-processed on a match-by-match basis with ImoClient software (Inmototec Object Tracking B.V., The Netherlands). Pre-processing included filtering with a weighted Gaussian algorithm (100% sensitivity), downsampling to 10 Hz, and automatic detection of possession and ball events based on synchronization of position tracking data with tagged event data.

Both tracking data and event data were imported in Python 3.8.3 and data processing and data analyses were conducted using the NumPy, Pandas, Math, SciPy, and Matplotlib libraries.

### 2.2. Success of Attacks

Every ball possession (attack) of each team was detected using the ball event data. A ball possession started with a team gaining control over the ball and ended whenever the opponent gained control over the ball again or there was a stoppage of play (ball out of bounds, free kick, corner, goal kick, or goal). To investigate passing behavior, the focus was on deliberate attacks only, as they include several deliberate (intentional) passes in a row. Therefore, only attacks that lasted longer than five seconds and had a minimum of three passes were selected.

The success of an attack was quantified with a danger value based on the tracking data. It was calculated with a zone value from which points were deducted depending on the defensive pressure of the opposing team. The zone value was measured using a grid similar to Link's work Dangerousity [14] which represents the last 35 m in front of the opposing goal, showing higher values the closer the player gets to the opposing goal. The amount of

defensive pressure was operationalized through the positions (regarding the position of the goal) and distances of the defenders in close proximity to the ball leading player.

For every passing reception, the danger value (ranging between 0 and 1) was calculated. Accordingly, attacks with a peak danger value  $>0$  were classified as successful and attacks with a peak danger value of 0 were classified as unsuccessful [28]. Hence, successful attacks represent a player in control over the ball, in at least 30 m range to the opposing goal with insufficient defensive pressure and the potential of creating a scoring opportunity.

### 2.3. Passing Effectiveness

Passing effectiveness was measured using an overall composite measure for the disruptiveness of the defensive organization following a pass, D-Def [15]. It is based on the idea that effective passes disrupt the opponents' defensive organization and create space to achieve scoring opportunities.

To measure defensive organization on the pitch, D-Def uses three different components: centroids (of full team, defensive-line, midfielder-line, and attacking-line), surface area of full team, and spread of full team. To measure the centroids of the formation lines, the tactical formations were automatically determined using a K-Means clustering ( $n\_clusters = 3$ ) algorithm of the players' average position in the first half. For every team, two superior formations, one attacking formation and one defending formation, were specified in three lines (e.g., 4-4-2, 4-3-3, 3-5-2). Accordingly, the players were assigned to the three lines.

The centroids of the full team ( $C_x, C_y$ ), of defensive-line ( $C_{xdef}, C_{ydef}$ ), of midfielder-line ( $C_{xmid}, C_{ymid}$ ), and of attacking-line ( $C_{xatt}, C_{yatt}$ ) were calculated by the average position in X- and Y-coordinates separately (Equations (1) and (2)). The surface area ( $S_{surface}$ ) was calculated as the smallest convex hull area of the positions of all players at a given timestamp  $t$  (Equation (3)). The spread ( $S_{spread}$ ) was calculated by the Frobenius norm of the positions of all players at a given timestamp  $t$  (Equation (3)). D-Def calculates the change of these three measures between the moment a pass was played and three seconds following this pass. According to the results of the principal component analysis in the work of Goes et al. [15], D-Def consists of three components (PC1, PC2, and PC3) (Equation (4)), which result in the measure of D-Def with a range from 0 to 150 (with 0 indicating no disruption and 150 indicating a maximum of disruption).

$$PC1 = -0.46 C_x + 0.26 C_y - 0.43 C_{xdef} + 0.24 C_{ydef} - 0.43 C_{xmid} + 0.24 C_{ymid} - 0.41 C_{xatt} + 0.24 C_{yatt} \quad (1)$$

$$PC2 = -0.26 C_x - 0.47 C_y - 0.24 C_{xdef} - 0.43 C_{ydef} - 0.25 C_{xmid} - 0.43 C_{ymid} - 0.24 C_{xatt} - 0.40 C_{yatt} \quad (2)$$

$$PC3 = 0.71 S_{area} + 0.71 S_{spread} \quad (3)$$

$$D - Def = |PC1| + |PC2| + |PC3| \quad (4)$$

For every attack that was considered, the last four passes were investigated. For every pass, the effectiveness (D-Def), the passing length ( $L_{pass}$  in [m]) (Equation (5)), passing velocity ( $V_{pass}$  in [m/s]) (Equation (6)), and passing angle ( $\alpha$  in [ $^\circ$ ]) were calculated (Equations (7) and (8)) (see Figure 1).

$$L_{pass} = \sqrt{(X_{reception} - X_{pass})^2 + (Y_{reception} - Y_{pass})^2} \quad (5)$$

$$V_{pass} = L_{pass} \div (t_{reception} - t_{pass}) \quad (6)$$

$$m_{pass} = (Y_{reception} - Y_{pass}) \div (X_{reception} - X_{pass}) \quad (7)$$

$$\alpha = \tan^{-1}(m_{pass}) \quad (8)$$

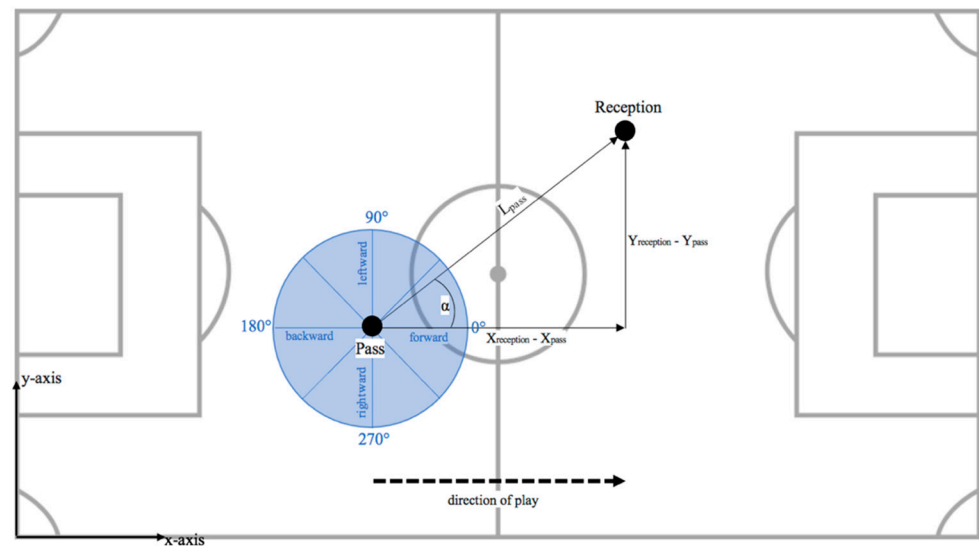
$m_{pass}$  = slope of passing vector

$\alpha$  = pass angle

$L_{pass}$  = pass length

$V_{pass}$  = pass velocity

$Y_{\text{reception}}$  = Y-coordinate of reception  
 $X_{\text{reception}}$  = X-coordinate of reception  
 $Y_{\text{pass}}$  = Y-coordinate of pass  
 $X_{\text{pass}}$  = Y-coordinate of pass  
 $t_{\text{reception}}$  = time of reception  
 $t_{\text{pass}}$  = time of pass



**Figure 1.** Graphical representation of passing length ( $L_{\text{pass}}$ ) with X- ( $X_{\text{reception}} - X_{\text{pass}}$ ) and Y-direction ( $Y_{\text{reception}} - Y_{\text{pass}}$ ) as well as the passing angle ( $\alpha$ ) of an exemplary pass from player 1 (Pass) to player 2 (Reception).

#### 2.4. Statistical Analyses

To test for the hypothesis that D-Def varies in successful and unsuccessful attacks, a two-way ANOVA ( $2 \times 4$ ) with repeated measures was conducted with repeated measure being the passing sequence (last 4 passes of an attack: pass 1 (last pass), pass 2 (penultimate pass), pass 3 (third last pass), and pass 4 (fourth last pass)) and success of an attack (groups: successful attacks vs unsuccessful attacks). If Mauchly's test of sphericity was significant the Greenhouse–Geisser correction was used.

To investigate the differences between successful and unsuccessful attacks in more detail and to account for the explanatory part of this study, independent *t*-tests were used to compare means of D-Def, pass length, and pass velocity of pass 1 (last pass), pass 2 (penultimate pass), pass 3 (third last pass), pass 4 (fourth last pass), average and maximum of an attack between successful and unsuccessful attacks. To avoid alpha-error, the Bonferroni–Holm correction was used. Before testing, Levene's test for homogeneity of variances was conducted for every comparison and if significant the correction was used.

To compare passing angles, circular statistics were used. A Watson–Williams F-test was conducted between successful and unsuccessful attacks for angles of every pass (pass 1, pass 2, pass 3, pass 4) separately. This test examines whether a set of mean directions are equal [29].

Additionally, effect sizes were calculated. In variance analyses partial eta squared ( $\eta^2$ ) was determined as effect size with  $\eta^2 < 0.06$  representing a small effect,  $0.06 \leq \eta^2 \leq 0.14$  representing a medium effect, and  $\eta^2 > 0.14$  representing a large effect [30]. For *t*-tests, Cohen's *d* was calculated as effect size with  $d < 0.5$  representing a small effect,  $0.5 \leq d \leq 0.8$  representing a medium effect, and  $d > 0.8$  representing a large effect [31].

All statistical analyses were conducted using IBM SPSS Statistics 25.0.0.0 (IBM Co., New York, NY, USA). The level of significance was set to  $p < 0.05$ .

### 3. Results

All 258 full matches included in this study resulted in 13,094 considered attacks. From all considered attacks, 7565 attacks were classified as unsuccessful (57.8%) and 5529 attacks were classified as successful (42.2%) with a mean *danger value* of  $0.27 \pm 0.26$ . Successful attacks lasted significantly longer (successful:  $40.41 \pm 74.42$  [s]; unsuccessful:  $29.20 \pm 71.00$  [s]) and had significantly more passes (successful:  $7.14 \pm 4.39$ ; unsuccessful:  $5.26 \pm 2.92$ ) compared to unsuccessful attacks ( $p < 0.001$ ).

The results of all ANOVAs are depicted in Table 1 and the results of pairwise comparisons and descriptive statistics between successful and unsuccessful attacks are shown in Table 2.

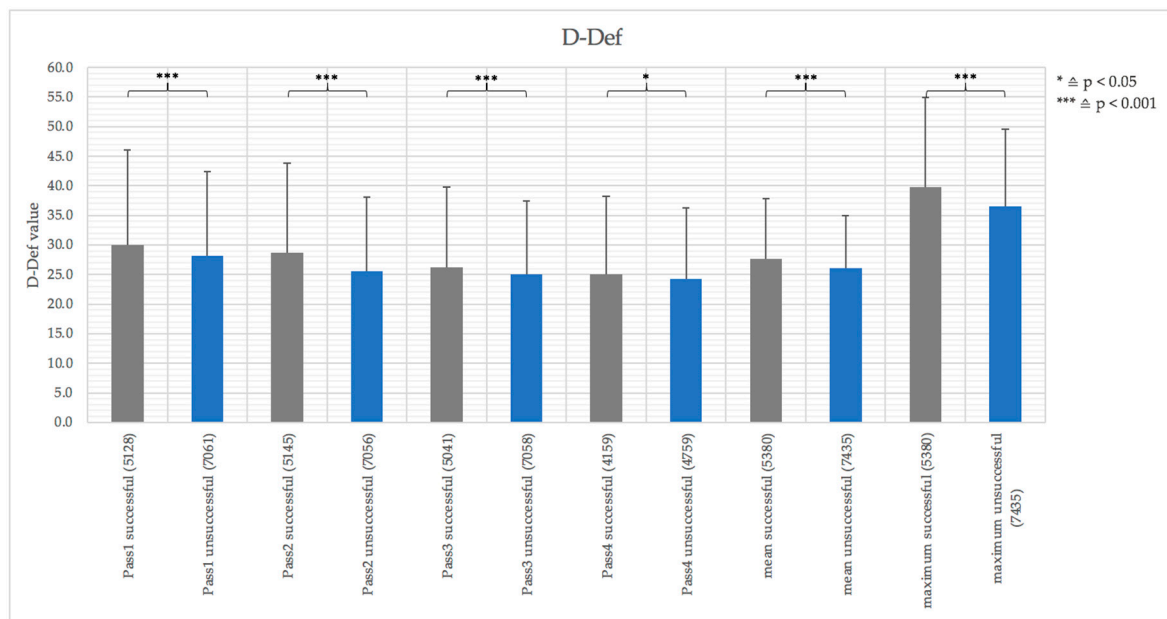
**Table 1.** Results of ANOVAs of D-Def, passing length, and passing velocity (\*  $\hat{=} p < 0.05$ ).

	df	Mean Square	F	p	$\eta^2$
D-Def					
Success of attacks (successful, unsuccessful attacks)	1	35,266.95	110.31	<0.001 *	0.14
Pass sequence (pass1, pass2, pass3, pass4)	3	24,909.81	185.04	<0.001 *	0.23
Success of attacks * Pass sequence error (success of attacks)	3	989.04	16.76	<0.001 *	<0.01
	8000	319.71			
Pass length					
Success of attacks (successful, unsuccessful attacks)	1	3658.50	22.47	<0.001 *	0.01
Pass sequence (pass1, pass2, pass3, pass4)	3	38,500.50	306.52	<0.001 *	0.12
Success of attacks * Pass sequence error (success of attacks)	3	852.80	6.79	<0.001 *	<0.01
	2258	162.85			
Pass velocity					
Success of attacks (successful, unsuccessful attacks)	1	604.96	16.38	<0.001 *	<0.01
Pass sequence (pass1, pass2, pass3, pass4)	3	14,718.74	497.91	<0.001 *	0.18
Success of attacks * Pass sequence error (success of attacks)	3	553.54	18.73	<0.001 *	<0.01
	2218	36.92			

**Table 2.** Results of *t*-tests between successful and unsuccessful attacks. Independent variables are depicted with D-Def values, passing length, and passing velocity of the last four passes of an attack as well as the mean and maximum of an attack.

	Successful Attacks			Unsuccessful Attacks			Independent <i>t</i> -Test			
	N	Mean	SD	N	Mean	SD	df	T	p	d
duration of attack (s)	5529	40.41	74.42	7565	29.20	71.00	13,092	8.75	<0.001	0.16
passes of attack	5529	7.14	4.39	7565	5.26	2.92	13,092	29.41	<0.001	0.52
D-Def pass 1	5128	29.96	16.02	7061	28.16	14.20	12,187	6.55	<0.001	0.12
D-Def pass 2	5145	28.69	15.13	7056	25.51	12.63	12,199	12.66	<0.001	0.23
D-Def pass 3	5041	26.26	13.56	7058	24.99	12.43	12,097	5.33	<0.001	0.10
D-Def pass 4	4159	25.01	13.21	4759	24.25	12.02	8916	2.87	0.029	0.06
D-Def mean	5380	27.67	10.10	7435	26.01	9.00	12,813	9.76	<0.001	0.18
D-Def max	5380	39.77	15.10	7435	36.54	13.03	12,813	12.95	<0.001	0.23
pass length pass 1 (m)	3441	23.97	15.93	4589	22.75	15.36	8028	3.44	0.005	0.08
pass length pass 2 (m)	3737	17.32	12.86	4691	16.19	11.48	8426	4.23	<0.001	0.09
pass length pass 3 (m)	3105	16.62	11.83	3259	16.28	11.03	6362	1.18	0.681	
pass length pass 4 (m)	2552	16.09	11.03	2238	16.47	10.82	4788	-1.21	0.681	
pass length mean (m)	5079	18.80	10.90	6573	17.74	10.71	11,650	5.26	<0.001	0.10
pass length max (m)	5079	26.86	15.55	6573	24.31	14.90	11,650	8.99	<0.001	0.17
pass velocity pass 1 (m/s)	3202	15.55	9.48	4192	16.41	10.37	7392	-3.66	0.002	-0.09
pass velocity pass 2 (m/s)	3711	10.50	4.49	4660	9.93	4.82	8369	5.58	<0.001	0.12
pass velocity pass 3 (m/s)	3085	10.37	4.61	3233	10.13	4.71	6316	2.05	0.245	
pass velocity pass 4 (m/s)	2541	10.12	4.39	2222	9.94	4.27	4761	1.44	0.600	
pass velocity mean (m/s)	4992	11.67	5.28	6362	11.87	6.50	11,352	-1.74	0.408	
pass velocity max (m/s)	4992	15.36	8.13	6362	15.42	9.14	11,352	-0.40	0.689	

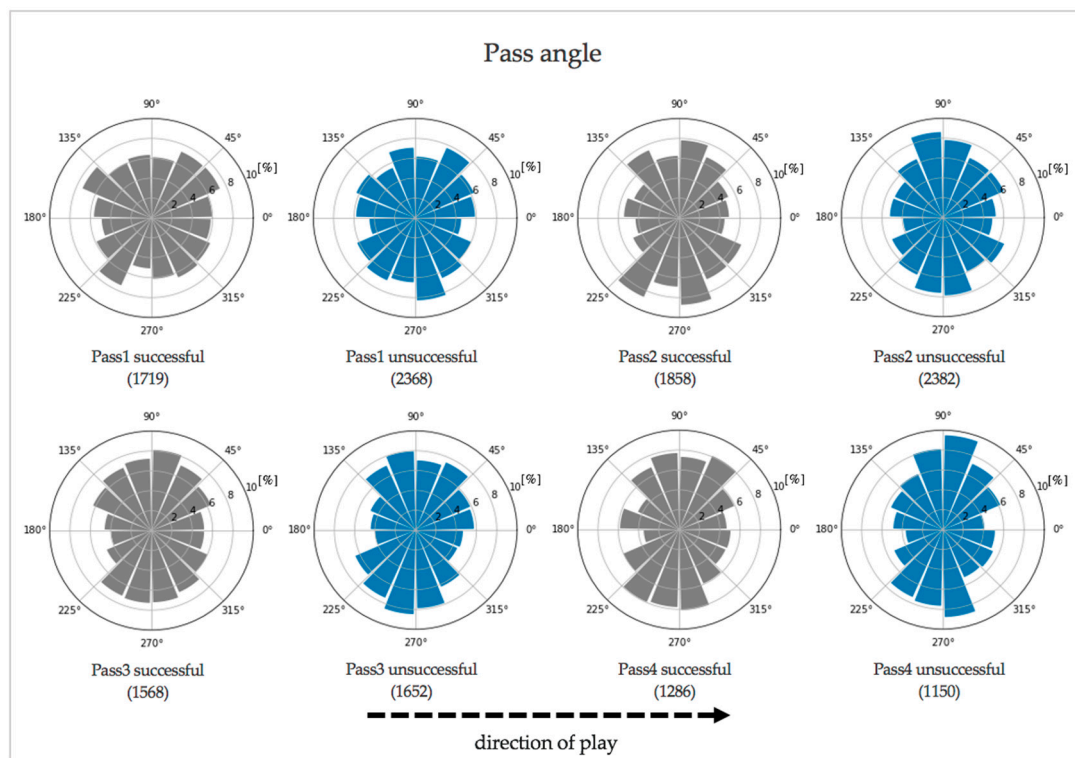
The ANOVA with D-Def as independent variable showed significant results between the passes in the attacking sequence ( $p < 0.001$ ), between successful and unsuccessful attacks, ( $p < 0.001$ ), and the interaction ( $p < 0.001$ ). While the interaction showed trivial effect size ( $\eta^2 < 0.01$ ) and the success effect showed medium effect size ( $\eta^2 = 0.14$ ), the passing sequence effect showed large effect size ( $\eta^2 = 0.23$ ). The pairwise comparisons of the individual passes ( $t$ -tests), depicted in Figure 2, indicated that D-Def values of all passes of successful attacks were significantly higher ( $0.001 < p \leq 0.029$ ) compared to unsuccessful attacks. Furthermore, the mean and maximum D-Def value of an attack was significantly higher in successful attacks compared to unsuccessful attacks ( $p < 0.001$ ) (see Figure 2). This difference showed the highest effect size in the penultimate pass ( $d = 0.23$ ) and the maximum of an attack ( $d = 0.23$ ).



**Figure 2.** Means and their standard deviation of D-Def values of the last four passes, as well as mean and maximum of these 4 last passes, separated in successful attacks (black) and unsuccessful attacks (blue) with associated sample sizes in brackets.

The ANOVAs with pass length and pass velocity as independent variable revealed similar results. The passing sequence, the success of an attack, and the interaction showed significant results in both variables ( $p < 0.001$ ). The difference between successful and unsuccessful attacks as well as the interaction showed trivial effect sizes ( $\eta^2 \leq 0.01$ ) for both passing length and velocity. The differences between the consecutive passes of an attack showed a medium effect size for passing length ( $\eta^2 = 0.12$ ) and a large effect size for passing velocity ( $\eta^2 = 0.18$ ). In the pairwise comparisons of individual passes between successful and unsuccessful attacks, pass length and pass velocity showed similar but inconsistent results. Pass 2 of successful attacks showed significantly longer and faster passes compared to unsuccessful attacks ( $p < 0.001$ ). Pass 3 and 4 showed no significant differences in passing length and velocity. Pass 1 of successful attacks showed smaller passing velocity compared to unsuccessful attacks ( $p = 0.002$ ).

The passing angles of the last four passes of successful attacks differed significantly from the passing angles of associated passes of unsuccessful attacks (Pass 1:  $p < 0.001$ ; Pass 2:  $p < 0.001$ ; Pass 3:  $p < 0.001$ ; Pass 4:  $p < 0.001$ ) (see Figure 3).



**Figure 3.** Percentage distribution of the passing angles of the last four passes of an attack grouped in successful (grey) and unsuccessful (blue) attacks with associated sample sizes in brackets.

#### 4. Discussion

The purpose of this study was to investigate passing in soccer using a seldomly used approach by investigating the consecutive chain of actions in the attacking process. Using this approach, we aimed (i) to show the validity of the pass effectiveness model, D-Def, and (ii) to reveal practice-relevant information about successful passing in elite soccer. To account for the first aim, we showed significant differences in D-Def of passes between successful and unsuccessful attacks. For the second purpose, we focus on the characteristics of passes of successful attacks and their differences to passes of unsuccessful attacks.

In general, the reported percentage distribution between successful (42.2%) and unsuccessful (57.8%) attacks showed fewer differences compared to studies using similar approaches. For example, Goes et al. [28] observed less than 10% successful attacks. This variation can be traced back to the less-strict inclusion criteria of a minimum of three consecutive passes to filter deliberate attacks compared to Goes et al. [28] who used the minimum of 5 s combined with a set starting location (in the first third) of an attack. The present approach increased the sample benefits and therefore, the statistical power and presentiveness of our analysis. Furthermore, the results of general attacking attributes (duration and number of passes) are comparable to results of investigations in other leagues [32]. Hence, the following results can be interpreted with a higher generality.

Assessing the effectiveness of passing, we used D-Def as a compositive measure of defensive disruptiveness following a pass. D-Def values consistently showed significant higher values in passes of successful attacks compared to passes of unsuccessful attacks (see Table 2), supporting the hypothesis that successful attacks are associated with passes with higher D-Def values. This result supports the idea that D-Def can be used to evaluated passing performance. This is in line with earlier studies demonstrating that D-Def can differentiate between good and bad passers on an individual level [15], and D-Def being a better predictor of the end result of a game in comparison to other key performance indicators [18]. This connection of D-Def with the end result can now be expanded with the current finding of D-Def showing higher values in successful attacks. This indicates the



connection of D-Def to success on a single possession level. Hence, it can be concluded, that D-Def is a valid measurement of passing effectiveness.

Furthermore, the analysis of consecutive actions in the chain of an attack reveals that this difference of D-Def values between successful and unsuccessful attacks is the biggest in the penultimate pass and the maximum of an attack. It can be derived, that one pass of an attack has to show a high D-Def value ( $D-Def > 28$ ) to result in a successful possession and that the penultimate pass (also called “hockey assist”) is a crucial moment of an attack in order to create time and space for the receiver to assist a scoring opportunity. This importance of the penultimate action (“hockey assist”) of an attack can be supported by the findings of Kempe et al. [26] who investigated the creativity of consecutive actions that led to goals. They discovered that the in goal scoring attacks the last three actions (hockey assist, assist, and shot on goal) of an attack are more creative than the previous ones and the creativity is increasing towards the end of an attack [26]. Comparably, a steady increase of D-Def values towards the last actions of an attack, with the last pass showing the highest D-Def values, was found in this study. Both results show that towards the end of an attack the increase of effectiveness of actions (creativity or passing effectiveness) is important for the success of an attack.

Subsequently, the results of the passing attributes (length, velocity, and angle) are discussed to answer the second research question of this paper about the characteristics of successful passing. The differences of passing length and passing velocity between successful and unsuccessful attacks showed trivial effect sizes in the ANOVAs and inconsistent results in *t*-tests. Those findings indicate that these two passing attributes are not highly decisive for successful passing. Furthermore, there are more prominent differences between the consecutive passes of an attack (medium effect size for passing length & large effect size for passing velocity in ANOVAs), for example, the last pass of an attack is strikingly longer and faster than the previous passes, compared to the differences between successful and unsuccessful attacks. In contrast, passing angles show significant differences between passes of successful and unsuccessful attacks. Overall, the proportions of passing angles reveal that few passes are played vertically forwards and backwards. In unsuccessful attacks, the distribution shows more sideways passes, and in successful attacks, more diagonal passes forward as well as backward. This outcome points out, that forward passes are not played very frequently and successful passing is not clearly connected to forward passing (as passes of successful attacks do not show more forward passes compared to passes of unsuccessful attacks). This strengthens the idea of the current approach that effective passing is not necessarily connected to forward passing which was a main criticism on other approaches evaluating passing behavior [10,12,33].

Summarizing those findings, passes of successful attacks are often played diagonally (forwards and backwards), with a medium length ( $\bar{L} = 18.8$  (m)), a medium velocity ( $\bar{V} = 11.7$  (m/s)) and a high D-Def value ( $\bar{D-Def} = 27.7$ ) indicating high disruption of the opposing defense. Those findings are comparable to earlier studies that discovered that effective passes are between 19–30 (m) long and with a passing velocity of 10.7 (m/s) [15]. However, the results of passing length and passing velocity should not be generalized as our results indicate that they are not highly decisive for successful passing. In contrast, for the success of an attack it is more important that on one side passes should not be played horizontally too frequently and on the other side, passes should show a high disruption of the opposing defense (high D-Def values). This disruption of the defense is mostly decisive as it creates space with fewer temporal and spatial pressure for receiving players and therefore can entail scoring opportunities in dangerous areas of the pitch. Those practice-relevant information about successful passing can help practitioners to create training regimes, analyze opponents, or rate players (e.g., for player recruitment), etc. [12]. Furthermore, the practice-oriented analysis of consecutive actions of an attack used in this study can help practitioners to transfer research findings to practice [2]. For example, the finding of the importance of the penultimate action of an attack can be implemented in training and match analyses.

There are some limitations of this study that should be noted. Firstly, the strong inclusion criteria to filter deliberate attacks compared to other approaches [28] result in a lot of short attacks that were excluded. Secondly, the high variance of the results due to the large sample size used should be noted in the interpretation of effect sizes. Lastly, in this investigation, only the last four passes of an attack were considered, and effects of previous passes are not depicted. However, both the findings of Kempe and Memmert [26] and this study indicate the importance of actions towards the end of an attack.

However, this investigation showed several strengths compared to similar studies. At first, the advanced study design makes it possible to reveal insights about the effectiveness of individual actions in the chain of an attack. Furthermore, the specific calculation of the investigated performance indicator D-Def takes the complex interactions with the opponent into account and therefore portrays the complex interaction of two soccer teams [34]. Additionally, by combining a large data set collected in professional competition (one full season of Dutch Eredivisie) with the methodology of D-Def that enables capturing complex patterns, conclusions can be made with a stronger ecological validity [4].

## 5. Conclusions

This study is an example of a sport science work that utilizes observational designs in which large data sets were collected in competition and used for the validation of a feature to assess some aspect of performance [3,12,14], in this case, D-Def [15]. We found that passes of successful attacks have significantly higher D-Def values compared to passes of unsuccessful attacks and thereby demonstrated the validity of this construct. Furthermore, we illustrated the constitutes of successful passing behavior to reveal practice-relevant information. Future research should focus on the collaboration between sport science and computer science, which will gain further importance to introduce the necessary expertise into the sport science domain [23] and will provide paths of new methods and key performance metrics into applied settings. This evolution of match analysis has the potential to change how soccer is analyzed and will define how a team can be successful in the future.

**Author Contributions:** Conceptualization, L.F. (Leander Forcher), M.K. and S.A.; methodology, L.F. (Leander Forcher), M.K. and S.A.; software, L.F. (Leander Forcher) and M.K.; validation, L.F. (Leander Forcher), M.K. and S.A.; formal analysis, L.F. (Leander Forcher); investigation, L.F. (Leander Forcher); resources, L.F. (Leander Forcher), M.K., S.A. and A.W.; data curation, L.F. (Leander Forcher) and M.K.; writing—original draft preparation, L.F. (Leander Forcher); writing—review and editing, M.K., S.A., L.F. (Leander Forcher), and A.W.; visualization, L.F. (Leander Forcher); supervision, M.K., S.A. and A.W.; project administration, A.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by grants of the Netherlands Organization for Scientific Research (629.004.012-SIA).

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of the University Medical Center Groningen (UMCG) (identification number 201800430, 1 November 2018).

**Informed Consent Statement:** Patient consent was waived due to the anonymization of all included tracking and event data. Therefore, no conclusions can be drawn about individual persons.

**Data Availability Statement:** Data examples are available upon request to the contact author.

**Acknowledgments:** This work was supported by a grant of the Netherlands Organization for Scientific Research (project title: The Secret of Playing Football: Brazil versus The Netherlands). Furthermore, we thank FC Groningen, Vitesse, and PSV Eindhoven as partners in this project. Additionally, we acknowledge support by the KIT-Publication Fund of the Karlsruhe Institute of Technology.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Herold, M.; Kempe, M.; Bauer, P.; Meyer, T. Attacking Key Performance Indicators in Soccer: Current Practice and Perceptions from the Elite to Youth Academy Level. *J. Sports Sci. Med.* **2021**, *20*, 158–169. [[CrossRef](#)]
2. Sarmiento, H.; Marcelino, R.; Campanico, J.; Matos, N.; Leitão, J. Match Analysis in Football: A Systematic Review. *J. Sports Sci.* **2014**, *32*, 1831–1843. [[CrossRef](#)] [[PubMed](#)]
3. Memmert, D.; Lemmink, K.A.P.M.; Sampaio, J. Current Approaches to Tactical Performance Analyses in Soccer Using Position Data. *Sports Med.* **2017**, *47*, 1–10. [[CrossRef](#)]
4. Goes, F.R.; Meerhoff, L.A.; Bueno, M.J.O.; Rodrigues, D.M.; Moura, F.A.; Brink, M.S.; Elferink-Gemser, M.T.; Knobbe, A.J.; Cunha, S.A.; Torres, R.S.; et al. Unlocking the Potential of Big Data to Support Tactical Performance Analysis in Professional Soccer: A Systematic Review. *Eur. J. Sport Sci.* **2021**, *21*, 481–496. [[CrossRef](#)] [[PubMed](#)]
5. Carling, C.; Bloomfield, J.; Nelson, L.; Reilly, T. The Role of Motion Analysis in Elite Soccer: Contemporary Performance Measurements Techniques and Work Rate Data. *Sports Med.* **2008**, *38*, 839–862. [[CrossRef](#)] [[PubMed](#)]
6. Moura, F.A.; Martins, L.E.B.; de O. Anido, R.; de Barros, R.M.L.; Cunha, S.A. Quantitative Analysis of Brazilian Football Players' Organisation on the Pitch. *Sports Biomech.* **2012**, *11*, 85–96. [[CrossRef](#)]
7. Clemente, F.; Afonso, J.; Sarmiento, H. Small-Sided Games: An Umbrella Review of Systematic Reviews and Meta-Analyses. *PLoS ONE* **2021**, *16*, e0247067. [[CrossRef](#)]
8. Sarmiento, H.; Clemente, F.M.; Araújo, D.; Davids, K.; McRobert, A.; Figueiredo, A. What Performance Analysts Need to Know About Research Trends in Association Football (2012–2016): A Systematic Review. *Sports Med.* **2018**, *48*, 799–836. [[CrossRef](#)]
9. Goes, F.; Schwarz, E.; Elferink-Gemser, M.; Lemmink, K.; Brink, M. A Risk-Reward Assessment of Passing Decisions: Comparison between Positional Roles Using Tracking Data from Professional Men's Soccer. *Sci. Med. Footb.* **2021**, 1–9. [[CrossRef](#)]
10. Power, P.; Ruiz, H.; Wei, X.; Lucey, P. Not All Passes Are Created Equal: Objectively Measuring the Risk and Reward of Passes in Soccer from Tracking Data. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 1605–1613.
11. Szczepański, Ł.; McHale, I. Beyond Completion Rate: Evaluating the Passing Ability of Footballers. *J. R. Stat. Soc. Ser. A* **2016**, *179*, 513–533. [[CrossRef](#)]
12. Rein, R.; Raabe, D.; Memmert, D. "Which Pass Is Better?" Novel Approaches to Assess Passing Effectiveness in Elite Soccer. *Hum. Mov. Sci.* **2017**, *55*, 172–181. [[CrossRef](#)]
13. Collet, C. The Possession Game? A Comparative Analysis of Ball Retention and Team Success in European and International Football, 2007–2010. *J. Sports Sci.* **2012**, *31*, 123–136. [[CrossRef](#)] [[PubMed](#)]
14. Link, D.; Lang, S.; Seidenschwarz, P. Real Time Quantification of Dangerousness in Football Using Spatiotemporal Tracking Data. *PLoS ONE* **2016**, *11*, e0168768. [[CrossRef](#)]
15. Goes, F.; Kempe, M.; Meerhoff, R.; Lemmink, K.A.P.M. Not Every Pass Can Be an Assist: A Data-Driven Model to Measure Pass Effectiveness in Professional Soccer Matches. *Big Data* **2018**, *7*, 57–70. [[CrossRef](#)]
16. Fernández, J.; Bornn, L.; Fraser, S. Wide Open Spaces: A Statistical Technique for Measuring Space Creation in Professional Soccer. In Proceedings of the MIT Sloan Sports Analytics Conference 2018, Boston, MA, USA, 23–24 February 2018; pp. 1–20.
17. Tenga, A.; Holme, I.; Ronglan, L.T.; Bahr, R. Effect of Playing Tactics on Goal Scoring in Norwegian Professional Soccer. *J. Sports Sci.* **2010**, *28*, 237–244. [[CrossRef](#)] [[PubMed](#)]
18. Kempe, M.; Goes, F. Move It or Lose It: Exploring the Relation of Defensive Disruptiveness and Team Success. Available online: <file:///C:/Users/MDPI/AppData/Local/Temp/EasyChair-Preprint-989.pdf> (accessed on 25 November 2021).
19. Lames, M. Chance Involvement in Goal Scoring in Football—an Empirical Approach. *Ger. J. Exerc. Sport Res.* **2018**, *48*. [[CrossRef](#)]
20. Tenga, A.; Holme, I.; Ronglan, L.T.; Bahr, R. Effect of Playing Tactics on Achieving Score-Box Possessions in a Random Series of Team Possessions from Norwegian Professional Soccer Matches. *J. Sports Sci.* **2010**, *28*, 245–255. [[CrossRef](#)]
21. Rathke, A. An Examination of Expected Goals and Shot Efficiency in Soccer. *J. Hum. Sport Exerc.* **2017**, *12*, 514–529. [[CrossRef](#)]
22. Lucey, P.; Bialkowski, A.; Monfort, M.; Carr, P.; Matthews, I. "Quality vs Quantity": Improved Shot Prediction in Soccer Using Strategic Features from Spatiotemporal Data. Available online: <https://www.sloansportsconference.com/research-papers/quality-vs-quantity-improved-shot-prediction-in-soccer-using-strategic-features-from-spatiotemporal-data> (accessed on 25 November 2021).
23. Memmert, D.; Rein, R. Match Analysis, Big Data and Tactics: Current Trends in Elite Soccer. *Dtsch. Z. Für Sportmed.* **2018**, *69*, 65–72. [[CrossRef](#)]
24. Ferrante, M.; Fonseca, G.; Carrari, A. A New Markovian Model for Tennis Matches. *Electron. J. Appl. Stat. Anal.* **2017**, *10*, 693–711. [[CrossRef](#)]
25. Goldner, K. A Markov Model of Football: Using Stochastic Processes to Model a Football Drive. *J. Quant. Anal. Sports* **2012**, *8*, 9. [[CrossRef](#)]
26. Kempe, M.; Memmert, D. "Good, Better, Creative": The Influence of Creativity on Goal Scoring in Elite Soccer. *J. Sports Sci.* **2018**, *36*, 2419–2423. [[CrossRef](#)]
27. Linke, D.; Link, D.; Lames, M. Football-Specific Validity of TRACAB's Optical Video Tracking Systems. *PLoS ONE* **2020**, *15*, e0230179. [[CrossRef](#)] [[PubMed](#)]

28. Goes, F.R.; Brink, M.S.; Elferink-Gemser, M.T.; Kempe, M.; Lemmink, K.A.P.M. The Tactics of Successful Attacks in Professional Association Football: Large-Scale Spatiotemporal Analysis of Dynamic Subgroups Using Position Tracking Data. *J. Sports Sci.* **2021**, *39*, 523–532. [[CrossRef](#)]
29. Watson, G.S.; Williams, E.J. On the Construction of Significance Tests on the Circle and the Sphere. *Biometrika* **1956**, *43*, 344–352. [[CrossRef](#)]
30. Field, A. *Discovering Statistics Using IBM SPSS Statistics*, 2nd ed.; SAGE: Los Angeles, CA, USA, 2005.
31. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Routledge: London, UK, 1988.
32. Tenga, A.; Sigmundstad, E. Characteristics of Goal-Scoring Possessions in Open Play: Comparing the Top, in-between and Bottom Teams from Professional Soccer League. *Int. J. Perform. Anal. Sport* **2011**, *11*, 545–552. [[CrossRef](#)]
33. Chawla, S.; Estephan, J.; Gudmundsson, J.; Horton, M. Classification of Passes in Football Matches Using Spatiotemporal Data. *ACM Trans. Spat. Algorithms Syst.* **2017**, *3*, 1–30. [[CrossRef](#)]
34. Grehaigne, J.-F.; Bouthier, D.; David, B. Dynamic-System Analysis of Opponent Relationships in Collective Actions in Soccer. *J. Sports Sci.* **1997**, *15*, 137–149. [[CrossRef](#)]