Scottish Journal of Political Economy

WILEY

SCOTTISH ECONOMIC SOCIETY

ORIGINAL ARTICLE OPEN ACCESS

Consensus, Diversity and the Wisdom of the Crowd: Predicting Final Rank Positions of the English Premier League

David O. Allen[1] | Don J. Webber[2]

[1]Independent Researcher, Bristol, UK | [2]Sheffield University Management School, The University of Sheffield, Sheffield, UK

Correspondence: Don J. Webber (d.j.webber@sheffield.ac.uk)

Received: 30 April 2024 | **Revised:** 13 April 2025 | **Accepted:** 15 April 2025

Funding: The authors received no specific funding for this work.

Keywords: betting | forecasting | predicting the unexpected | soccer

ABSTRACT

This study assesses the Wisdom of the Crowd premise using a natural experiment. We collected primary data relating to predictions of the final football club rank positions in the English Premier League over two consecutive seasons using an innovative game and compared these predictions with predictions made by sporting pundits, football blogs and inferred from betting odds. Our results question the Wisdom of the Crowd premise, and we show that scores generated by the Diversity Prediction Theorem vary from one season to the next, suggesting that crowds are not adept at predicting the unexpected.

JEL Classification: L83, D01, Z29

1 | Introduction

Forecasts are an everyday occurrence in a wide variety of life events, with some predictions crucial for everyday functioning and others associated with leisure activities and pastimes. Predictions tend to associate with people's behaviours and preferences (Granger and Pesaran 2000) rather than purely rational information processing. The collective opinion of a group of independent and diverse individuals is referred to as the 'wisdom of the crowd', and groups of sports fans may generate more accurate predictions of end of season rank positions than individual experts. Evaluations of forecasts of football (soccer) matches is not uncommon in academic journal articles (e.g., Reade et al. 2021, among others), though evaluations of the wisdom of the crowd in making predictions of sporting events remain rare.

This paper offers a simple methodology for quantifying and comparing predictions of final rank positions of clubs in a competition which we then use to test whether the wisdom of the crowd premise holds. We constructed an easily manageable game whereby our 'players' predicted the end of season rank positions of all 20 football clubs within the English Premier League (EPL) where the winner is the player who predicts the final club rank positions across the whole EPL most accurately. We source players from four sets: we ran the game in two public houses located in the English city of Bristol[1] and captured public house customers' predictions in 'pub A' and 'pub B' for the seasons 2018/19 and 2019/20; we compared predictions by the crowds in pubs A and B with two sets of predictions provided by professional sports correspondents, which are 'the pundits' and the aggregate odds offered online by betting organisations, that is, 'the bookies', for both EPL seasons.

While the initial motivation for creating this game was lighthearted, our game offers the opportunity to conduct a natural experiment and evaluate the central premise of Galton (1907) and Surowiecki (2004) that groups of people can make better predictions than an individual or a group of experts, otherwise known as the 'wisdom of the crowd' phenomenon. The evidence for crowd wisdom generally comes from experiments were

people guess a specific value to a defined problem; in the case of Galton (1907), this was a country fair competition where people were invited to guess the weight of a butchered ox. The general outcome was that while no individual guessed the correct amount, the average of all the guesses is typically the nearest to the true weight. These experiments, and many others (e.g., guess the number of jellybeans in a jar), generally produce the same results, but as a phenomenon it is not fully explained or understood.

Page (2007a) offered what he saw as proof of the wisdom of the crowd phenomenon with his Diversity Prediction Theorem (DPT). An important aspect of Page's (2007a) theorem is the role played by diversity, specifically where diversity is the spread in the approach to seeing a problem (or cognitive diversity) rather than diversity based on socioeconomic categories (i.e., age, ethnicity, etc.). In Page (2007a, 2007b) the problem's setting is institutional and the problem to solve is likely to be a complex organisational goal—something more than guessing the number of jelly beans in a jar. The form of Page's (2007a) theorem is:

$$CE \equiv AE - DIV \tag{1}$$

where *CE* is the crowd error (i.e., the squared difference between the crowd's average prediction, and the actual value), *AE* is the average error (i.e., the total of the squared differences between the crowd's predictions and the actual value), and *DIV* is the Diversity (i.e., the variance in the crowd's predictions). Given the identity, *CE* must be less than or equal to *AE* with the implication that for *CE* to be small (i.e., to prove the crowd got it right), *AE* needs to be large; and since *AE* is large, *DIV* must be large to make *CE* small.

In the absence of any detailed algebra, Page (2007b) explains the terms of the DPT using a worked example based on the predictions of seven 'prognosticators' for the top 12 picks for the 2005 NFL in the USA (Page 2007b, 12–13). In our worked example (Table 1) we also use seven players who are predicting the top six ranks in an imaginary football league. Panel 1 presents the end of season rank, *r*, for the top six clubs and the rank predicted, *p*, by the players (A–G), along with the average of the seven predictions for each club, $\bar{p}$. The Crowd Error (*CE*), $\sum(r-\bar{p})^2$, is 27.31. Diversity (*DIV*) is the total variance in the predictions and the value here is 14.41. In panel 2, the average error (*AE*), $\sum(r-p)^2$, is 41.71 and thus the crowd's prediction is more accurate than its average member (i.e., $CE < AE$).

Page (2007b) does not elaborate on how the *CE* can be interpreted other than the inference that when $CE < AE$ the crowd is less wrong than the average player. For example, the *CE* for Liverpool is 13.80 (panel 1) which compares to the *AE* of 15.43 (panel 2).

The error on a per-rank basis (EPR) can be estimated by taking the square root of the average of the value of interest; for example, the *CE* of 27.13 divided by 6 players' predictions is 4.55, and it's square root is ±2.13. This suggests that on

TABLE 1 | Worked example of Page (2007) Diversity Prediction Theorem.

		Player's predictions (*p*)								Diff	CE	DIV
Panel 1 club	**Rank (*r*)**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**($\bar{p}$)**	**($r-\bar{p}$)**	**($r-\bar{p}$)²**	**VAR**
Liverpool	1	5	5	2	6	4	5	6	4.71	−3.71	13.80	1.63
Man City	2	1	3	5	3	5	6	4	3.86	−1.86	3.45	2.41
Spurs	3	2	2	6	2	1	2	2	2.43	0.57	0.33	2.24
Man Utd	4	6	1	1	1	6	3	5	3.29	0.71	0.51	4.78
Arsenal	5	3	6	3	4	3	1	1	3.00	2.00	4.00	2.57
Chelsea	6	4	4	4	5	2	4	3	3.71	2.29	5.22	0.78
		46	32	42	54	35	38	46	43.63	Total	27.31	14.41

		($r-p$)²								
Panel 2 Club	**Rank (*r*)**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**AE ($r-p$)²**	
Liverpool	1	16	16	1	25	9	16	25	15.43	$CE \equiv AE - DIV$
Man City	2	1	1	9	1	9	16	4	5.86	$27.31 \equiv 41.71 - 14.41$
Spurs	3	1	1	9	1	4	1	1	2.57	*CE*: Crowd error
Man Utd	4	4	9	9	9	4	1	1	5.29	*AE*: Average error
Arsenal	5	4	1	4	1	4	16	16	6.57	*DIV*: Diversity
Chelsea	6	4	4	4	1	16	4	9	6.00	
Total AE		30	32	36	38	46	54	56	41.71	

Note: Players are demoted 'A' to 'G'; in this respect, player 'A' has the lowest Total AE at 30—their predictions are closest to the actual league outcomes.

average the crowd prediction was out by ±2.13 ranks. For the average player ($AE = 41.71$) the prediction was out by ±2.64 ranks; for player a, their predictions were out by ±2.24 ranks, and so on.

This DPT is not without its critics, and it attracted immediate criticism where some dismissed it as a mathematical tautology (Yudkowsky 2007). More recently, Nobre and Fontanari (2020) contend that guesses are not normally distributed—neither are ours in our analysis below—which explains why some individuals' guesses are better than a *CE* score, and hence the crowd does not always arrive at the best answer. Another issue relates to the skewness in the predictions, and this is a reflection of reality often informed by socialised expectations but also in our case because of the rank nature of the game, where, for example, a dominance of predictions of a club in rank position 20 will have a distribution tail that tends towards higher rank positions only. Nevertheless, the theorem and the concept have caught the imagination of the broader public, particularly within the business community who were quick to share the ideas of Surowiecki (2004) and Page (2007a, 2007b) and to promote the idea of cognitive diversity.

This paper provides three main contributions to the literature. First, it highlights that the practical applications of the DPT or the wisdom of the crowd are, at best, limited to single event scenarios. This is because games based on single estimates (e.g., predicting the number of jellybeans in a jar) readily allow for a square root of the *CE* when estimating differences between a consensus guess and the actual number of jellybeans in the jar, but this cannot be done as easily with games involving multiple guesses or rank choices. In our game, we had a mixture of both consensus and diversity within the same total diversity score. Second, our empirical evidence reveals that experts (i.e., the bookies and the pundits) can be outsmarted by talented (or lucky) individual member(s) of a crowd, though this is not evidence that supports the central premise of the wisdom of the crowd. Third, we present evidence consistent with Page's (2007a) claim that the wisdom of the crowd is not strong at predicting unexpected outcomes, which raises questions about the broader usefulness of the concept.

This paper contains six sections. The next section explains the mechanics of the game and its scoring and draws attention to the similarities between our game and Page's (2007a) DPT. Section 3 profiles the groups used in the experiment (two pubs, the pundits and the bookies). Section 4 explores the depiction and measurement of internal agreement and consensus within groups of players and examines the parameters underpinning the diversity scores (DIV). Section 5 presents the results which compare the groups' predictions of the EPL final league rank positions to those inferred from the odds offered by the bookies, and we show why AEs and crowd errors can fluctuate between seasons. Section 6 concludes.

2 | Our Game

Interest in the global sport of association football (soccer) has been high since before 1863 when its rules were codified in England thousands of years after it was played in its original forms. The economics literature has retained a keen focus on the game since Sloane's (1971) seminal paper on the football club as a utility maximiser. Contributions to the literature include examinations of multi-argument utility functions within football clubs (Madden and Robinson 2012), investments in player talent when a football club is more interested in value maximisation than in profit maximisation (Prinz and Thiem 2021), and assessments of the competitive balance in the European Champions League (Triguero-Ruiz and Avila-Cano 2023). In this paper we assess the ability to correctly predict the final English Premiership League (EPL) rank positions in a given season, where the player who most accurately predicts the final rank positions of all 20 clubs wins the game.

2.1 | Participation and Scoring

Individual players participating in our game make a prediction of the final rank positions[2] of all EPL clubs, i. This type of information can be captured using a survey, though we tailored our approach to the collection of predictions from our players as a game to create an element of interpersonal competition and to strengthen belonging between friends and acquaintances within a social pub setting. The score for each player is calculated by taking the difference (or error) between the prediction rank of a particular club in the league table, p, and that club's actual rank, r, at the end of the season. For example, if a player predicts Chelsea will finish in 5th but at the end of the season Chelsea finish 10th, then the error is −5. This error is squared to give the 'difference score' $\left((p_i - r_i)^2\right)$ for that player's club prediction (e.g., $-5 \times -5 = 25$). Difference scores for each club are added up to give the player's Total Score (TS), where $TS = \sum_{i=1}^{20} (p_i - r_i)^2$. The smaller the difference score then the closer a player's predictions are to the true final league rank positions. It follows that if a player correctly predicts the final rank positions of all EPL clubs then their total score will equal zero. A natural component of this approach is that if a player's prediction for only one club is very inaccurate, such as it being wayward by 15 rank positions, while all other predictions are only one rank out, then that player will have a TS of $[(19 \times 1) + (15 \times 15) = 19 + 225 =] 244$. As there is often a surprise performance by a club, such as Leicester City's difficult to predict performance in 2015/16 when they won the EPL, this surprise performance adds to the banter within the pub atmosphere.

While our game predates Page's (2007a) DPT, as we first ran this game in 1997, there are similarities in that the difference score calculated in our game is the same as that used to derive the *AE* in the DPT. Note that the best score in our game (zero) does not compare to examples of estimating a single value which is under/over-estimated (e.g., the number of jellybeans in a jar, pages in a book or, the weight of an ox), and therefore the errors that make up our *AE* are not normally distributed.

Our game has three types of players: pub players, pundits, and bookies. We collected data across these three categories of players for the 2018/19 and 2019/20 EPL seasons. The combined numbers of players in pubs A and B were 33 and 39 for these two seasons, respectively.

2.2 | Pub Players

The profile of our pub players was consistent with what one would expect to find in an average English public house: typically white, male and aged 30 years or over. In each year of running our game, there were a number of female players, though we do not claim that our sample is representative of any population. While many of our players were knowledgeable of the EPL, there were always some with little football knowledge who appeared to enter either to join in with a group or to participate for fun. Most players chose to complete the entry form at home, and their predictions are made without recognisable peer group pressure, peer discussion or peer cooperation, where we refer to peers as other players in this game. The players were informed that the game was based on the premise that once predictions were submitted to the gamemaster, then they could not be altered. Anecdotal observations suggested that pub customers sometimes struggled to form an impartial prediction when it came to their favourite (and sometimes least favourite) EPL club. Pub groups in years with more players showed more diverse knowledge of football.

2.3 | Pundits

Pundit predictions were sourced from websites before the start of each season. Full sets of rank predictions for all 20 clubs were difficult to source, as many pundits appeared prepared only to predict in batches (e.g., top six and bottom three). Given the probabilities of making correct rank predictions for all 20 clubs, many pundits might have felt such a task was a fool's errand. Nevertheless, there are quite a few pundits on social media (e.g., YouTube and Reddit) who were prepared to make a full set of end-of-season rank predictions for all 20 clubs.

We would expect a pundit's knowledge of football matters to be more comprehensive than an average pub player's knowledge of the game. The full list of pundits in our sample for 2018/19 included Andy Edwards (NBC), Daily Star Fans, Dan Karell (NBC), ESPN, FiveThirtyEight, Guardian, Joe Prince-Wright (NBC), Kyle Bonn (NBC), Matt Reed (NBC), Nicholas Mendola (NBC), Phil McNulty (BBC), and the Sky Supercomputer, whereas the full list of pundits in our sample for 2019–20 was BT—BIG DATA, Every Tip, Guardian, Last Word, Michael Owen, Phil McNulty (BBC Sports), Sport Nation. Bet, TJS-101, and the Radio Times. We do not claim that either set of pundits is representative of all contemporary sporting correspondents nor of all blogs dedicated to football.

2.4 | Bookies

Bookmakers use oddsmakers that evaluate each club's likely performance in the coming season and, like the average pub player or pundit, bookmakers assess a club's past performance to predict their final rank position in the forthcoming season. An oddsmaker's initial assessment sets the true odds or probabilities such that they add up to 100%. Bookmakers then adjust these odds to achieve a mark-up (e.g., 110%) and they adjust the odds to reflect typical betting patterns which usually see more bets placed on leading clubs (e.g., Manchester City, Manchester Utd, Liverpool, Spurs, Arsenal, and Chelsea—commonly referred to as 'the big six'). These odds can change as the start of the season nears, where changes could be the result of changes in betting patterns (to balance the books) or new information about a club (e.g., new manager or success in the transfer market).

The odds offered by the bookies are rarely on a specific league rank position (e.g., 11th), they propose more general bets such as the chances of a club winning the EPL, the chances of being in the top four, or the chances of being relegated. The odds that we used related to the chance of winning the EPL. A main feature of these published odds is that many are batched and hence more than one club can share the same odds of winning the league. It is also the case that the same odds are offered across different bookmakers for a given club, which is counter to the notion or spirit of diversity in the context of the DPT. Table 2 lists the average odds (in rank order) used in the 2018/19 and 2019/20 seasons sourced after the end of the previous season. In the context of this game, bookies are the experts, as they have a strong incentive to understand what is happening across the EPL.[3]

TABLE 2 | Average percentage odds of winning the PFL.

Rank (GP)	2018/19	Ave	2019/20	Ave
1	Man City	57.74	Man City	62.28
2	Liverpool	25.36	Liverpool	30.55
3	Man Utd	13.63	Spurs	5.75
4	Chelsea	7.93	Chelsea	3.64
5	Spurs	6.57	Man Utd	3.37
6	Arsenal	4.51	Arsenal	2.22
7	Wolves	0.47	Wolves	0.62
8	Everton	0.44	Everton	0.56
9	Leicester	0.35	Leicester	0.43
10	West Ham	0.33	Newcastle	0.29
11	Southampton	0.20	West Ham	0.22
12	Newcastle	0.17	Watford	0.13
13	Crystal Palace	0.16	Southampton	0.12
14	Burnley	0.14	Crystal Palace	0.11
15	Fulham	0.12	Aston Villa	0.11
16	Watford	0.12	Bournemouth	0.10
17	Bournemouth	0.11	Brighton	0.08
18	Brighton	0.10	Burnley	0.08
19	Huddersfield	0.09	Sheffield Utd	0.07
20	Cardiff	0.08	Norwich	0.07
Total		118.65	Total	110.80
$n =$		23		27

Note: n is the number of bookies estimates used in that year.

3 | Data

Once all pub players' entry forms were submitted the group place (*GP*) can be calculated. Note that descriptions of all variables are included at the end of the script. For the Bookies the process starts by converting all fractional odds into percentages, and the averages of these odds is then ranked to give the 20 rank positions for all EPL clubs, as in Table 1. While this *GP* calculation is crude it is at least loosely based on some initial assessment or interpretation of a club's recent relative performance and an indication of the predicted rank before the start of the season. For the pub teams and the Pundits, the GP is the rank of the average of the predicted ranks per club, that is, $\overline{p}$ in Table 1. By ranking this average, a club is either in, say, 2nd or 3rd and not in 2.37. As an easily understood integer, the *GP* allows players to evaluate their individual rank predictions relative to those made by the rest of their group when that information is subsequently made available and again, for the group, it is an indicator of the likely end-of-season rank position as predicted before the start of the season.

Table 2 shows the predictions made by the pundits for the 2018/19 season, and it reveals both the group averages per club and the GP based on the rank of these averages. Table 2 shows that the group average for Man Utd and Chelsea was equal at 4.58. Where two clubs have the same average predicted rank, then the club that finished in a lower rank in the previous season is given the lower rank, including newly-promoted clubs; for this reason, Chelsea's GP was manually adjusted to 5th.

The relationship between the mean value and the GP is not regular but weaves either side of *GP*, and the six plots in Figure 1 show that this occurs for all groups to differing degrees, although there are similarities whereby the mean is generally higher than the rank positions from 1st to 5th and below the 14th rank.

3.1 | Conditions for Diversity to Prevail Over Ability

Page (2007b, 9) identified four conditions that must be met for 'diverse groups, on average, to outperform groups of the best individuals.' In the context of our game these conditions should apply to the two pub teams, as we treat the pundits and the bookies as the 'groups of best individuals'. Page (2007b, 10) summarises these four conditions as (a) the problem must be difficult to predict, (b) the people must be smart, (c) the people must be diverse and (d) the groups must be reasonably big and chosen from a large pool. Our pub teams met the first three of these conditions and it is the fourth that might be difficult to achieve as it remains unclear whether what is 'reasonably big'.

4 | Diversity or Consensus Within a Group

When making a judgement, individual estimates and predictions tend to be subject to an element of idiosyncratic bias. When individuals are independent, Page's (2007a) DPT states that averaging the estimates of a large crowd eliminates this bias, and therefore crowds are believed to make better judgements when compared with an expert. Specifically, Page's (2007a, 2007b) DPT is where the squared error of a crowd's collective prediction is equal to the average squared error minus the predictive diversity; when the diversity within a crowd is large then the error of the crowd is small, and thus the crowd makes better estimates and predictions than does an individual expert.

The DPT theorem is built on several foundations. First, socio-economic tests of the DPT tend to be based on single shot issues such as guessing the weight of a butchered ox at a country fair (Galton 1907) or the modern fair version, which is guessing the number of jellybeans in a jar. Although few, if any, individuals guess the specific amount correctly, the mean of all guesses across a diverse crowd is typically nearer to the true amount than would be a guess of an individual expert. However, few studies test the strength of the DPT by replacing the estimate of a single amount with estimates of a set of interconnected complex predictions. Our game described above allows us to test whether the DPT remains strong when there is a need to predict a set of interconnected issues (i.e., final league rank positions of football clubs).

Second, the DPT is based on a crowd competing against an expert, but few economic studies pool a group of experts together as experts. Using a collection of experts could be useful when making important predictions, as this could reduce bias in their individual predictions. Although experts tend to be very knowledgeable about something (or very skilful in a particular area), they may be affected by other types of bias, such as peer pressure and the need to illustrate conformity to each other and/or to other expectations (Table 3).

Third, the DPT is based on the assumption that individuals within a crowd (or indeed across an expert group) are independent.[4] True independence may not exist in the real world due to cultural upbringings, religious beliefs and social pressures to conform to (or outdo) a social group. Even the predictions of a single expert are likely to be at least in part informed and influenced by their interactions with other experts, just like pub customers may be influenced by other pub customers through spatial autocorrelation and spatial clustering of latent effects. In what follows we test Page's (2007a) DTP using the game outlined above and using the following test statistics.

4.1 | Diversity Scores (DIV)

We estimated prediction diversity at the start of each season by calculating total *DIV* of the predictions for each club. Without the *AE* as a reference, the total *DIV* score is not easily interpreted, so we plotted the individual club *DIV* scores to reveal relative differences in predictions across clubs. Figure 2 plots these *DIV* scores when ranked by the *GP* for pubs A and B and for the pundits. In 2018/19, pub A's *DIV* score for Wolves indicates a high level of diversity in the group's predictions for this club. In the same year, there was a lack of agreement for Southampton and Wolves in pub B, and for West Ham, Wolves, Crystal Palace, and Newcastle for the pundits. Patterns across the six panels highlight low diversity of predictions across the first six rank positions, with greater diversity in mid-table, and then generally lower diversity for rank positions 17th to 20th, and this resonates with Figure 1.

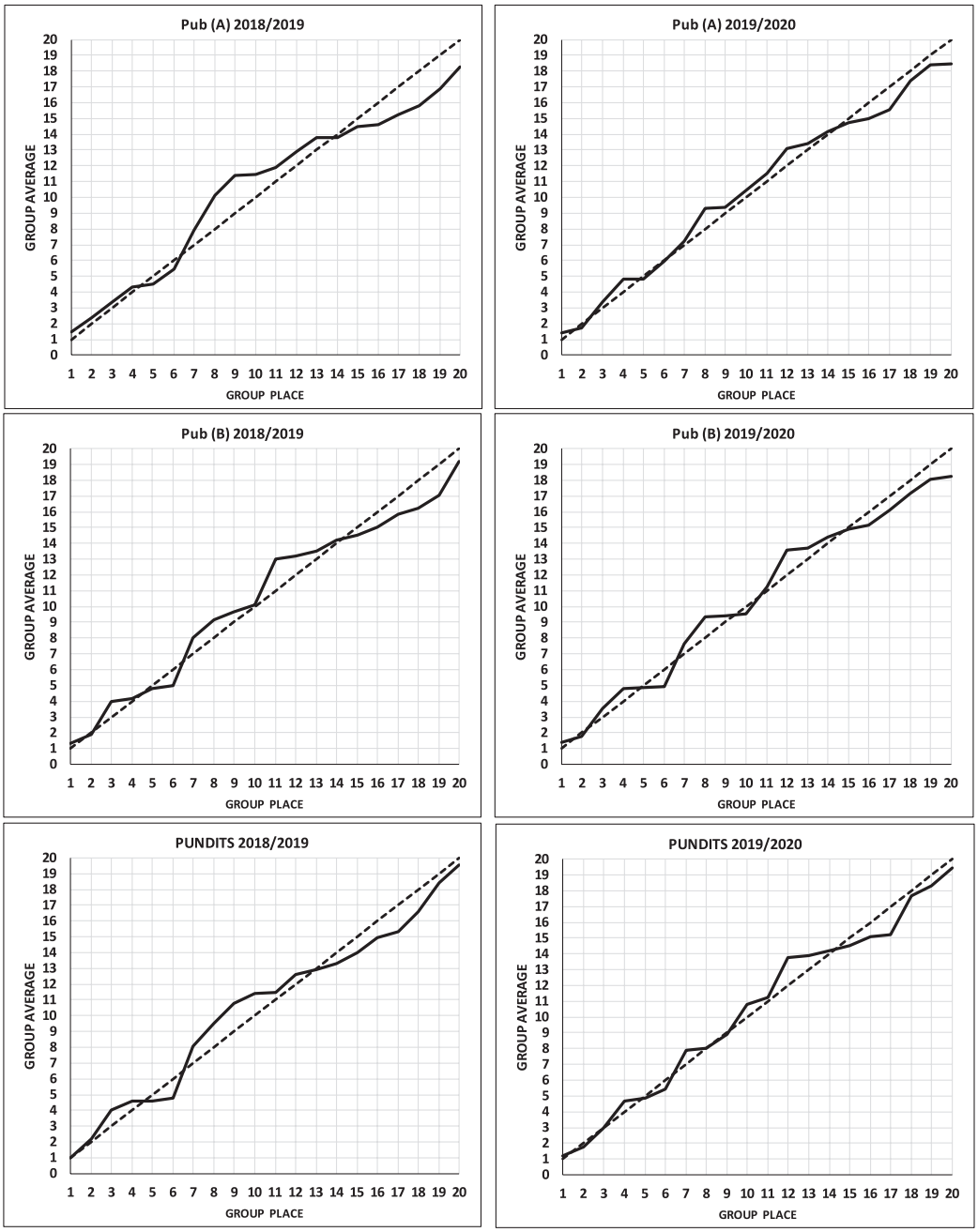

FIGURE 1 | Group average and group place. The number of observations vary with Pub A having 20 (24), Pub B having 13 (15), and the Pundits having 12 (9) in year 2018/19 (2019/20) respectively. The dashed line reflects the group place, and the solid line reflects the group average.

In terms of the DPT, we have a mixture of both consensus and diversity within the same total diversity score. Consensus and diversity are determined by the number of different predictions and the range in the predicted rank positions for each club. To see the relationship between these two parameters requires a different way of setting out the prediction data.

4.2 | Prediction Matrices

Tables 4 and 5 show the number of predicted rank positions for each club over the 2018/19 and 2019/20 football seasons batched by pub A, pub B, and the pundits. These tables can be read by column (league rank position 1–20) or row (football club rank position) and are symmetrical in that the sum of the rows and the sum of the columns add up to the number of players in a group.

4.3 | *X*-Scores (Predictions Per Club) and *Z*-Scores (Predictions Per League Rank Position)

In Table 5, for pub A, 12 players predicted Man City would finish 1st, seven predicted 2nd, and one predicted 3rd, which add up to three different and unique ranks predicted by our 20 players. Column *X* gives the total number of different or unique

TABLE 3 | Tabulated predictions: Pundits 2018/19.

Pundits—Predict the league—2018/2019 season														
Premier League team	**Andy Edwards [NBC]**	**Joe Prince-Wright [NBC]**	**Guardian**	**Matt Reed [NBC]**	**Phil McNulty [BBC]**	**Nicholas Mendola [NBC]**	**Dan Karell [NBC]**	**Five thirty eight**	**Kyle Bonn [NBC]**	**ESPN**	**Daily Star fans**	**Sky super computer**	**Group place (GP)**	**Group ave**
Man City	1	1	1	1	1	1	1	1	1	1	1	1	1	1.00
Liverpool	2	2	2	2	2	3	2	2	2	2	2	3	2	2.17
Spurs	4	4	4	6	4	6	4	3	4	3	4	2	3	4.00
Man Utd	6	3	3	5	3	2	7	6	6	4	6	4	4	4.58
Chelsea	5	6	6	3	5	4	5	4	3	6	3	5	5	4.58
Arsenal	3	5	5	4	6	5	3	5	5	5	5	6	6	4.75
Everton	8	7	7	7	8	7	6	10	7	10	10	10	7	8.08
Leicester	9	10	9	11	10	9	11	8	8	14	8	7	8	9.50
West Ham	7	8	8	8	9	8	9	13	13	11	18	17	9	10.75
Wolves	12	9	10	10	7	13	10	17	10	12	7	20	10	11.42
Fulham	14	12	15	9	11	11	12	15	9	9	9	12	11	11.50
Crystal Palace	10	15	16	12	13	10	18	7	18	8	15	9	12	12.58
Burnley	13	16	11	13	12	16	8	14	16	15	13	8	13	12.92
Southampton	15	11	12	18	16	15	13	9	12	13	12	14	14	13.33
Newcastle	17	14	13	15	18	12	15	11	15	7	20	11	15	14.00
Bournemouth	11	13	17	14	14	17	14	16	14	17	19	13	16	14.92
Brighton	18	17	14	16	15	14	16	18	11	16	14	15	17	15.33
Watford	16	18	18	17	19	18	17	12	19	18	11	16	18	16.58
Huddersfield	19	19	19	19	17	19	19	19	17	19	16	19	19	18.42
Cardiff City	20	20	20	20	20	20	20	20	20	20	17	18	20	19.58

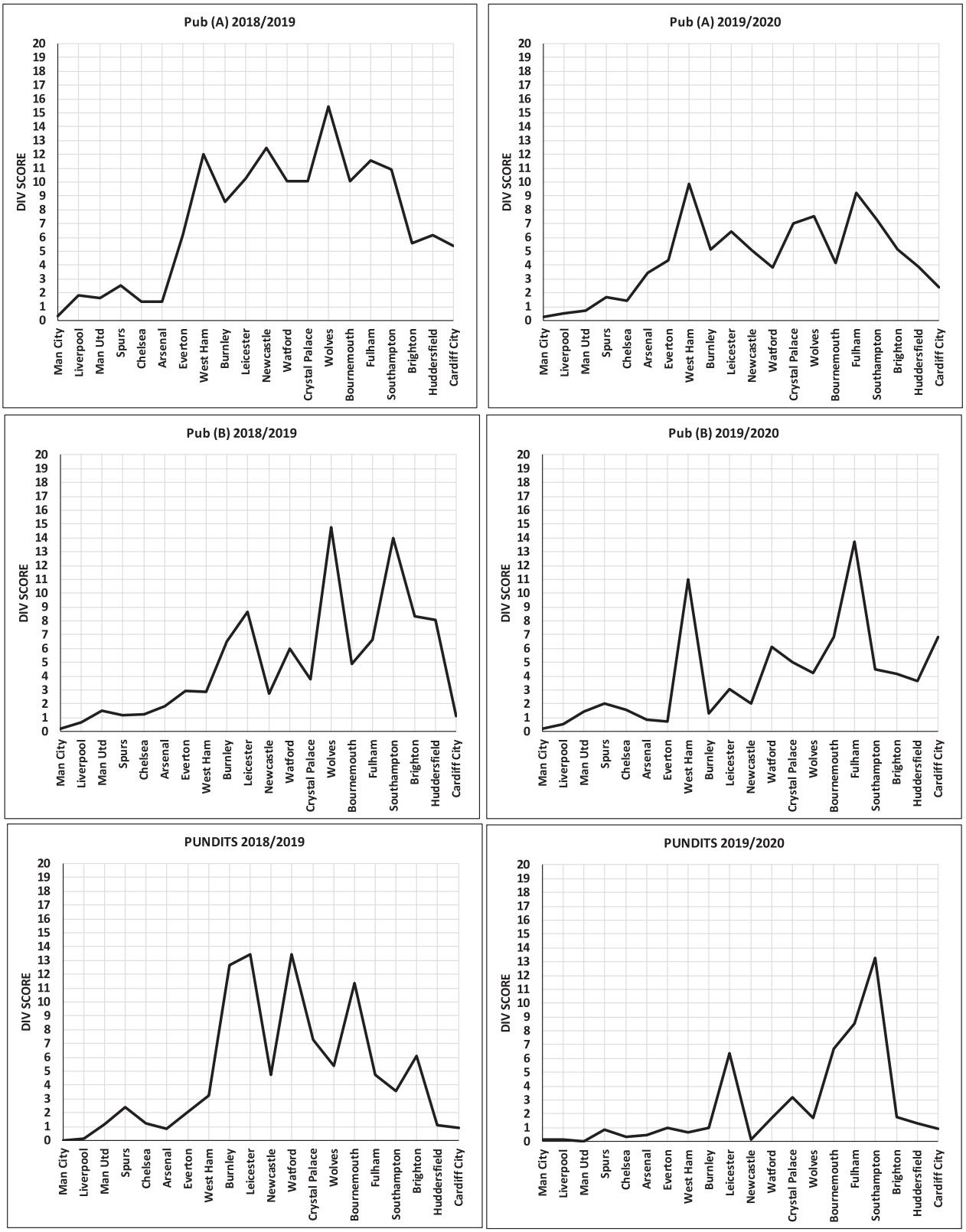

FIGURE 2 | Diversity (DIV) scores for pub A, pub B, and the pundits. In each case the horizontal axis shows the final order of the teams in the season and the vertical axis shows the diversity scores (DIV).

predicted rank positions for each club, which for Man City was 3, and this column reflects the *X*-score for each club estimated by the players from pub A. For pub A, in 2018/19 there were 171 unique predictions for the 20 clubs.

Looking down the columns of Table 4 shows that 12 players predicted that the champions would be Man City, five thought Liverpool, two went for Man Utd, and one chose Spurs. Row *Z* gives the count of possible champions as 4, which is our

TABLE 4 | Prediction matrices, 2018/19.

	GP	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	*X*	DIV
Pub (A) 2018/19																							
Man City	1	12	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0.35
Liverpool	2	5	10	1	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	1.83
Man Utd	3	2	2	7	5	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	1.64
Spurs	4	1	1	4	6	3	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	7	2.51
Chelsea	5	0	0	5	5	6	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.35
Arsenal	6	0	0	2	1	6	9	1	1	0	0	0	0	0	0	0	0	0	0	0	0	6	1.35
Everton	7	0	0	0	0	0	2	11	4	1	1	0	0	0	0	0	0	0	1	0	0	6	6.19
West Ham	8	0	0	0	0	0	0	4	5	4	1	1	1	1	0	0	1	1	0	1	0	10	11.99
Burnley	9	0	0	0	0	0	0	0	2	4	5	2	1	1	1	2	1	0	0	1	0	10	8.54
Leicester	10	0	0	0	1	0	0	2	1	1	2	1	5	1	3	1	1	1	0	0	0	12	10.25
Newcastle	11	0	0	0	0	1	0	0	2	4	2	1	0	3	2	1	2	0	2	0	0	10	12.49
Watford	12	0	0	0	0	0	0	0	1	2	3	2	2	3	1	0	1	3	2	0	0	10	10.09
Crystal Palace	13	0	0	0	0	0	0	0	1	0	0	5	3	1	3	2	1	0	1	2	1	10	10.06
Wolves	14	0	0	0	0	0	0	1	1	2	1	2	2	0	1	2	1	2	3	1	1	13	15.46
Bournemouth	15	0	0	0	0	0	0	0	1	0	1	3	1	1	3	1	2	4	1	1	1	12	10.05
Fulham	16	0	0	0	0	0	0	0	0	1	2	2	1	3	0	3	1	3	0	2	2	10	11.54
Southampton	17	0	0	0	0	0	0	0	0	1	1	0	2	3	3	1	1	2	1	2	3	11	10.89
Brighton	18	0	0	0	0	0	0	0	0	0	1	1	0	2	0	2	6	2	5	1	0	8	5.56
Huddersfield	19	0	0	0	0	0	0	0	0	0	0	0	1	1	2	4	1	0	2	7	2	8	6.19
Cardiff City	20	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	2	2	2	10	8	5.39
	Z	4	4	6	6	6	6	6	11	9	11	10	11	11	10	11	13	9	10	10	7	8.6	143.71

(Continues)

TABLE 4 | (Continued)

	GP	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	*X*	DIV
Pub (B) 2018/19																							
Man City	1	9	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.21
Liverpool	2	4	7	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0.69
Man Utd	3	0	2	2	5	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.54
Chelsea	4	0	0	5	3	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1.21
Arsenal	5	0	0	2	4	2	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1.25
Spurs	6	0	0	3	0	6	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	4	1.85
Everton	7	0	0	0	0	0	1	7	1	2	0	1	1	0	0	0	0	0	0	0	0	6	2.92
Burnley	8	0	0	0	0	0	0	2	4	3	0	2	2	0	0	0	0	0	0	0	0	5	2.90
West Ham	9	0	0	0	0	0	0	1	4	4	1	0	2	0	0	0	0	1	0	0	0	6	6.52
Leicester	10	0	0	0	0	0	0	3	2	1	2	3	0	0	0	1	0	1	0	0	0	7	8.69
Crystal Palace	11	0	0	0	0	0	0	0	0	0	0	3	2	4	2	1	0	1	0	0	0	6	2.77
Newcastle	12	0	0	0	0	0	0	0	0	1	2	0	1	4	0	3	0	2	0	0	0	6	6.02
Bournemouth	13	0	0	0	0	0	0	0	0	0	1	1	3	1	2	2	3	0	0	0	0	7	3.79
Southampton	14	0	0	0	0	0	0	0	1	0	2	2	1	0		1	0	2	3	0	1	8	14.79
Fulham	15	0	0	0	0	0	0	0	0	0	2	0	0	0	3	4	1	3	0	0	0	5	4.86
Watford	16	0	0	0	0	0	0	0	0	1	0	0	1	2	1	0	5	0	3	0	0	6	6.62
Wolves	17	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1		3	2	2	9	13.98
Brighton	18	0	0	0	0	0	0	0	0	0	1	0	0	1	3	0	1	1	2	3	1	8	8.33
Huddersfield	19	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0	2	0	2	4	2	6	8.07
Cardiff City	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	4	7	3	1.10
	Z	2	3	5	4	4	5	4	6	7	9	7	8	6	6	7	6	8	5	4	5	5.6	98.11

(Continues)

TABLE 4 | (Continued)

	GP	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	*X*	DIV
Pundits 2018/19																							
Man City	1	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.00
Liverpool	2	0	10	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.14
Spurs	3	0	1	2	7	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1.17
Man Utd	4	0	1	3	2	1	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	6	2.41
Chelsea	5	0	0	3	2	4	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1.24
Arsenal	6	0	0	2	1	7	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0.85
Everton	7	0	0	0	0	0	1	5	2	0	4	0	0	0	0	0	0	0	0	0	0	4	2.08
Leicester	8	0	0	0	0	0	0	1	3	3	2	2	0	0	1	0	0	0	0	0	0	6	3.25
West Ham	9	0	0	0	0	0	0	1	4	2	0	1	0	2	0	0	0	1	1	0	0	7	12.69
Wolves	10	0	0	0	0	0	0	2	0	1	4	0	2	1	0	0	0	1	0	0	1	7	13.41
Fulham	11	0	0	0	0	0	0	0	0	4	0	2	3	0	1	2	0	0	0	0	0	5	4.75
Crystal Palace	12	0	0	0	0	0	0	1	1	1	2	0	1	1	0	2	1	0	2	0	0	9	13.41
Burnley	13	0	0	0	0	0	0	0	2	0	0	1	1	3	1	1	3	0	0	0	0	7	7.24
Southampton	14	0	0	0	0	0	0	0	0	1	0	1	3	2	1	2	1	0	1	0	0	8	5.39
Newcastle	15	0	0	0	0	0	0	1	0	0	0	2	1	1	1	3	0	1	1	0	1	9	11.33
Bournemouth	16	0	0	0	0	0	0	0	0	0	0	1	0	2	4	0	1	3	0	1	0	6	4.74
Brighton	17	0	0	0	0	0	0	0	0	0	0	1	0	0	3	2	3	1	2	0	0	6	3.56
Watford	18	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	2	2	4	2	0	6	6.08
Huddersfield	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	9	0	3	1.08
Cardiff City	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	10	3	0.91
	Z	1	3	5	4	3	5	7	5	6	4	9	7	7	7	6	7	8	7	3	3	5.4	95.72

Note: Z is the number of different clubs for each league rank position, *X* is the total number of different predicted rank positions for each club, and DIV is the diversity score for each club.

TABLE 5 | Prediction matrices, 2019/20.

	GP	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	*X*	DIV
Pub (A) 2019/20																							
Man City	1	14	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.24
Liverpool	2	10	12	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0.54
Spurs	3	0	1	17	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0.73
Man Utd	4	0	0	4	7	6	5	1	1	0	0	0	0	0	0	0	0	0	0	0	0	6	1.66
Arsenal	5	0	1	2	7	6	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	6	1.41
Chelsea	6	0	0	0	4	8	7	1	1	2	0	0	1	0	0	0	0	0	0	0	0	7	3.46
Everton	7	0	0	0	1	2	2	13	5	0	0	0	0	0	0	0	1	0	0	0	0	6	4.33
Wolves	8	0	0	0	1	0	1	3	5	8	3	0	0	0	0	1	1	0	0	1	0	9	9.87
Leicester	9	0	0	0	0	0	0	3	9	4	4	0	0	1	2	1	0	0	0	0	0	7	5.15
West Ham	10	0	0	0	0	0	1	2	2	1	7	6	3	0	1	0	0	0	0	1	0	9	6.41
Watford	11	0	0	0	0	0	0	0	0	3	7	4	4	3	1	0	0	1	1	0	0	8	5.08
Bournemouth	12	0	0	0	0	0	0	0	0	1	0	5	3	6	3	4	1	0	1	0	0	8	3.83
Crystal Palace	13	0	0	0	0	0	0	0	1	1	2	2	2	4	3	4	2	1	2	0	0	11	6.99
Burnley	14	0	0	0	0	0	0	0	0	3	0	1	2	2	5	3	2	4	1	1	0	10	7.56
Aston Villa	15	0	0	0	0	0	0	0	0	0	0	2	3	2	2	5	6	2	2	0	0	8	4.12
Newcastle	16	0	0	0	0	0	0	0	0	0	1	2	3	4	2	1	3	2	2	1	3	11	9.25
Southampton	17	0	0	0	0	0	0	0	0	0	0	2	2	1	4	4	2	3	2	1	3	10	7.25
Brighton	18	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	3	3	6	8	1	8	5.15
Sheffield Utd	19	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	6	1	3	11	6	3.90
Norwich	20	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	2	6	8	6	6	2.41
	Z	2	4	4	7	5	7	7	7	9	6	8	10	9	10	9	11	9	10	8	5	7.4	89.36

(Continues)

TABLE 5 | (Continued)

	GP	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	*X*	DIV
Pub (B) 2019/20																							
Man City	1	9	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.24
Liverpool	2	6	6	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0.56
Spurs	3	0	3	5	5	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	1.45
Chelsea	4	0	0	4	2	4	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	5	2.03
Arsenal	5	0	0	2	5	3	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.58
Man Utd	6	0	0	1	3	8	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0.86
Everton	7	0	0	0	0	0	1	6	5	3	0	0	0	0	0	0	0	0	0	0	0	4	0.76
Wolves	8	0	0	0	0	0	3	3	3	0	1	2	0	1	0	1	0	1	0	0	0	8	11.02
West Ham	9	0	0	0	0	0	0	1	2	4	7	0	1	0	0	0	0	0	0	0	0	5	1.31
Leicester	10	0	0	0	0	0	0	1	4	4	2	2	0	2	0	0	0	0	0	0	0	6	3.05
Watford	11	0	0	0	0	0	0	1	0	0	2	4	6	2	0	0	0	0	0	0	0	5	2.06
Crystal Palace	12	0	0	0	0	0	0	0	0	0	1	1	4	4	1	0	2	0	1	1	0	8	6.11
Bournemouth	13	0	0	0	0	0	0	0	0	0	2	0	1	5	3	1	0	2	1	0	0	7	5.00
Burnley	14	0	0	0	0	0	0	0	0	1	0	0	1	1	5	3	2	1	1	0	0	8	4.24
Southampton	15	0	0	0	0	0	0	0	0	0	0	3	1	0	2	2	2	2	2	1	0	8	6.86
Newcastle	16	0	0	0	0	0	0	0	0	3	0	1	0	0	1	1	1	3	3	1	1	9	13.72
Aston Villa	17	0	0	0	0	0	0	0	0	0	0	0	1	0	2	4	3		1	4	0	6	4.52
Brighton	18	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	4	5	2	1	7	4.16
Norwich	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	1	0	4	5	5	3.66
Sheffield Utd	20	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	1	2	8	7	6.86
	Z	2	3	5	4	3	6	7	5	5	6	8	7	6	7	8	8	8	8	7	4	5.9	80.04

(Continues)

TABLE 5 | (Continued)

	GP	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	*X*	DIV
Pundits 2019/20																							
Man City	1	7	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.17
Liverpool	2	2	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.17
Spurs	3	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.00
Chelsea	4	0	0	0	6	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0.89
Man Utd	5	0	0	0	2	6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0.32
Arsenal	6	0	0	0	1	3	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0.47
Leicester	7	0	0	0	0	0	0	4	3	1	1	0	0	0	0	0	0	0	0	0	0	4	0.99
Everton	8	0	0	0	0	0	0	3	3	3	0	0	0	0	0	0	0	0	0	0	0	3	0.67
Wolves	9	0	0	0	0	0	0	1	2	3	3	0	0	0	0	0	0	0	0	0	0	4	0.99
West Ham	10	0	0	0	0	0	0	1	0	2	2	2	0	0	0	2	0	0	0	0	0	5	6.40
Watford	11	0	0	0	0	0	0	0	0	0	0	7	2	0	0	0	0	0	0	0	0	2	0.17
Southampton	12	0	0	0	0	0	0	0	0	0	0	0	1	3	4	0	0	1	0	0	0	4	1.73
Bournemouth	13	0	0	0	0	0	0	0	0	0	0	0	2	3	1	2	0	0	1	0	0	5	3.21
Aston Villa	14	0	0	0	0	0	0	0	0	0	0	0	1	2	2	2	2	0	0	0	0	5	1.73
Crystal Palace	15	0	0	0	0	0	0	0	0	0	2	0	0	0	1	1	3	2	0	0	0	5	6.69
Burnley	16	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	0	1	2	1	0	8	8.54
Newcastle	17	0	0	0	0	0	0	0	1	0	0	0	2	0	0	1	1	0	3	0	1	6	13.28
Brighton	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	1	2	1	5	1.78
Norwich	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	2	4	1	5	1.33
Sheffield Utd	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	6	3	0.91
	Z	2	2	1	3	2	3	4	4	4	5	2	6	4	5	6	5	6	5	4	4	3.9	50.44

Note: Z is the number of different clubs for each league rank position, *X* is the total number of different predicted rank positions for each club, and DIV is the diversity score for each club. The shading highlights that greater diversity occurs at lower ranked teams.

Z-score for 1st rank position, and the total of these *Z* scores is also 171.

In comparison, pub B's predictions saw only two potential contenders for champions: Man City (9) and Liverpool (4), while the 12 pundits were unanimous in their belief that only Man City would be crowned champions in 2018/19 (which turned out to be correct).

4.4 | Line of Complete Agreement (CA)

Cell values along the diagonal from coordinates 1:1 to 20:20 indicate the degree of complete agreement (*CA*). In the case of the 12 pundits in the 2018/19 season (Table 4), if they were unanimous in all their predictions, then the value in each of the 20 cells across the diagonal would be 12 and the sum of all these predictions would be 240 (12×20). Likewise, all the values in column *X* and row *Z* would be 1, and their summations would each be 12.

Table 6 presents the summary statistics from the prediction matrices for pubs A and B and reveals that the distributions of *X* and *Z* scores are generally symmetrical. The *DIV* scores are mixed, as evidenced from the six plots in Figure 2. The matrices, along with the *X* and *Z* scores, offer a way to compare predictions for individual clubs across groups. For example, in Tables 4 and 5, the *DIV* scores for Wolves in the Pub A and Pub B predictions are high and are explained by the higher number of unique predictions and the higher range of the predictions made over the two seasons. The Pearson's correlations in Table 6 confirm the strong relationship between the *X* scores and *DIV*.

4.5 | Two Ratios of Consensus (or Lack of Diversity)

Tables 4 and 5 show that there is consensus in predictions for some teams (the 'Big Six' in particular), suggesting the need for summary measures of consensus that relate to all the predictions (as opposed to club specific measures such as the *X* score). The Parameters column in Table 6 lists the number of potential predictions (*PP*) a group can make when playing the game and the count of actual unique predictions made (*UP*). This property enabled us to calculate two summary statistics to capture the degree of consensus across players within a group. These measures and their calculations are listed in Table 7, along with the DIV from Tables 4 and 5.

First, we calculated the ratio of unique predictions (*UP*) to potential predictions (*PP*), expressed as a percentage, which we call *Ratio #1*. Here, a lower value indicates more consensus. For example, if pub A had CA on all club rank positions in 2018/19, then the number of unique predictions would be 20, which would give an actual to potential ratio of 5%.

We calculated the ratio of the sum of predictions along the line of *CA* to the sum of all predictions, which we call *Ratio #2*. The sum of CA is the sum of the diagonal cells in the prediction matrices in Tables 4 and 5. For example, for pub A in 2018/19, cell 1:1 equals 12 and cell 2:2 equals 10, etc.; these values are totalled up to give 103, which equates to 25.8% of all predictions made. Here, a higher value indicates greater consensus.

These two ratios corroborate Tables 4 and 5 and Figures 1 and 2, which show that the pundits have more consensus in predictions, whereas pub A has the greatest diversity.

4.6 | Comparing Group Predictions (GP)

Comparisons of predictions made by the groups can be identified from the matrices. Table 8 presents *GP* listed in accordance with the rank positions predicted by the bookies' *GP*. Each column presents the club's *X*-score from each group while the total number of players in each group is provided in square brackets. Note that because the individual bookies use the same odds for more than one club so the *X*- and *Z*-scores are not symmetrical; therefore, any prediction matrix for the bookies would be misleading, and likewise, we cannot calculate an overall *DIV* score or a *DIV* score for each club. We can generate a proxy for the *X*-score based on the number of different odds placed on a given club. The clubs selected for the top six rank positions are predictable, and the *X*-scores relative to the group size suggest a higher degree of consensus. Similarly, those predicted to be relegated typically include two newly promoted clubs, but here there is less consensus. It is in the middle of the table that there is a marked lack of consensus, and this is particularly the case for predictions made by pub teams, although the bookies were very unsure about Sheffield Utd's chances in 2019/20.

Table 8 presents us with a conundrum in terms of the likely AE once the EPL final rank positions are known: if the consensus for the top 6 and, say, bottom two holds true, then how do we explain the role being played by diversity in determining the AE?

5 | Output and Results

During and at the end of the season, players are provided with two output sheets which show how each player's predictions are performing or has performed. Tables 9 and 10 present examples from pub B at the end of the 2019/2020 season where clubs are listed in their final rank position. Difference scores (SCORE) and rankings (RANKING) were published as in Table 8, with individual difference scores per team (Table 9) published less frequently (every 2 months). By ranking each players' score (low to high) we can pick out the median score (e.g., 420 for individual 'D' in Table 8). The ticks and crosses to the right of a club indicate whether the group's *GP* predictions are close to the actual rank. This is calculated by taking the total of the difference scores for each club (Table 9) and ranking them to find the top and bottom three clubs.

5.1 | Group Score

The group score is calculated in the same way as the player's difference scores, i.e., the actual rank minus the group rank, all squared. Recall that *GP* is the rank of the averages of predictions.

TABLE 6 | Summary statistics for prediction matrices.

	Skew	Kurtosis	Parameters		Correlations	
Pub (A) 2018/19						
X—scores (By Club)	−0.24	−0.65	*n* =	20		
Range (per club)	−0.59	−0.85	*n*×20	400	X/DIV	0.918
Z-scores (by League position)	−0.29	−1.16	Predictions (Count)	171	Range/DIV	0.920
DIV score	−0.05	−1.23	S Diagonal (CA)	103		
Pub (A) 2019/20						
X—scores (By Club)	−0.31	0.08	*n* =	24		
Range (per club)	−0.18	0.08	*n*×20	480	X/DIV	0.890
Z-scores (by League position)	−0.52	−0.36	Predictions (Count)	147	Range/DIV	0.853
DIV score	0.23	−0.69	S Diagonal	134		
Pub (B) 2018/19						
X—scores (By Club)	0.00	−0.18	*n* =	13		
Range (per club)	0.04	−1.22	*n*×20	260	X/DIV	0.859
Z-scores (by League position)	0.03	−0.37	Predictions (Count)	111	Range/DIV	0.930
DIV score	1.02	0.65	S Diagonal (CA)	70		
Pub (B) 2019/20						
X—scores (By Club)	−0.17	−0.59	*n* =	15		
Range (per club)	−0.02	−0.88	*n*×20	300	X/DIV	0.846
Z-scores (by League position)	−0.49	−0.82	Predictions (Count)	117	Range/DIV	0.879
DIV score	1.31	1.96	S Diagonal (CA)	80		
Pundits 2018/19						
X—scores (By Club)	−0.13	−0.54	*n* =	12		
Range (per club)	0.04	0.00	*n*×20	240	X/DIV	0.831
Z-scores (by League position)	−0.27	−0.64	Predictions (Count)	107	Range/DIV	0.936
DIV score	0.90	−0.35	S Diagonal (CA)	81		
Pundits 2019/20						
X—scores (By Club)	0.44	0.35	*n* =	9		
Range (per club)	1.07	1.04	*n*×20	180	X/DIV	0.747
Z-scores (by League position)	−0.22	−0.87	Predictions (Count)	77	Range/DIV	0.954
DIV score	1.87	3.86	S Diagonal (CA)	77		

Note: *n* is the number of players in the group, *X* score is the total number of unique predicted rank positions per club, *Z* score is the number of different clubs per rank position (or place), PP is the number of potential predictions (*n*×20), UP is the number of unique predictions ($\sum$ *X*-scores, or *Z*-scores), and CA is the diagonal on a prediction matrix.

This group score is similar to but not the same as the *AE* used in Page's (2007a) DPT, which would be the average of all scores (i.e., 402.4).

Table 9 shows that player I won the 2019/2020 competition in pub B with a score of 224, which compares to the Group Score of 394 (and an *AE* of 402.4). Table 9 also illustrates the point that a winning score does not necessarily depend on the number of exact predictions (see player B, 3rd). Table 10 shows how each player's score and the group score add up. Here the worksheet is formatted to highlight individual difference scores equal to zero (i.e., an exact prediction) and those greater than 100 (i.e., predictions that are out by 10 or more ranks).

For the 2019/2020 season, players in pub B as a group greatly underestimated Sheffield Utd's performance as they predicted this club to finish in 20th when it actually finished in 9th, which is a difference score of (11×11 =) 121. It is worth pointing out that pub B's *GP* difference score of 121 for Sheffield Utd represented just over 30% of their overall group score of 394.

TABLE 7 | Summary statistics for internal agreement/disagreement.

	n	PP	UP	% UP	APC	ΣCA	% CA	DIV
Pub (A) 2018/19	20	400	171	42.8%	8.6	103	25.8%	143.7
Pub (A) 2019/20	24	480	147	30.6%	7.4	134	27.9%	89.4
Pub (B) 2018/19	13	260	111	42.7%	5.6	70	26.9%	98.1
Pub (B) 2019/20	15	300	117	39.0%	5.9	80	26.7%	80.0
Pundits 2018/19	12	240	107	44.6%	5.4	81	33.8%	95.7
Pundits 2019/20	9	180	77	42.8%	3.9	77	42.8%	50.4
Ratios of consensus				*#1*			*#2*	

Note: PP is the number of potential predictions ($n \times 20$), UP is the number of unique predictions ($\sum X$-scores, or Z-scores), % UP is the number of unique predictions (UP) as a percentage of the number of potential predictions (PP), APC is the average number of rank predictions per club (UP ÷ 20), and CA is the diagonal on a prediction matrix.

5.2 | Results Across the Four Groups

Table 11 lists the top three scores, the highest (or worst) score, and the group score for all groups of players. It includes the values that make up Page's (2007a) DPT ($CE = AE - DIV$) and the second and third agreement ratios (APC and CA%). Comparison of players, pundits, and bookies on a like-for-like basis requires us to use each group's *AE* which also means that we can compare individual difference scores to that of the bookies. We investigate the results from three perspectives: (i) Page's (2007a) DPT and the wisdom of the crowd, (ii) whether 'experts' are better able to predict EPL rank positions compared to pub customers and (iii) the results across the two seasons.

5.2.1 | Page's DPT and the Wisdom of the Crowd

Table 11 reveals that in all cases the pub teams and the pundits saw lower crowd error (*CE*) than the *AE*, indicating that the crowd is more accurate than its average member. This corroborates Page's (2007a) DPT. Page's (2007b), 13) 'brilliance of the crowds' occurs when the *CE* is lower than any of the individual estimates, although there is only one instance where this is the case: in 2018/19, pub A scored a *CE* of 114.2 and an *AE* of 126. Excluding the results for pub A in the 2018/19 season, there were 5 occasions when *CE* scores feature in the top 3 results (highlighted using arrows in Table 11).[5] Over the 2 years the *CE* score is always lower than the *AE* score and yet it is only in 2018/19 that the *CE* would make it into the top 4 when it came 2nd.

It is difficult to conclude that the concept of the wisdom of the crowd has prevailed given that there are individuals who have performed better than the crowd in the majority of the games. Perhaps more of a concern is that we cannot identify predictions made by the crowd that generated this lower score. This could be important if guesses/predictions relate to policy options or election candidates. In games based on single estimates (e.g., predicting the number of jellybeans in a jar) we can take the square root of *CE* to estimate the difference between the consensus guess and the actual number, but this becomes more difficult with games involving multiple guesses or choices. Thus, practical applications of DPT or the wisdom of the crowd are, at best, limited to single event scenarios.

5.2.2 | Comparing Pub Teams to the Pundits and Bookies

Comparison of the groups in 2018/19 and 2019/20 shows that the pundits' *AE* score was lower than the corresponding *AE* scores for the two pubs, yet the bookies' *AE* score was lower still, as shown in Table 12. By this metric, the consensus among the experts seems to have won the day, as they showed greater ability than the pub teams when it came to predicting final EPL rank positions.

Notwithstanding this, inspection of the top three in the 2018/19 season reveals a reversed picture in that 1st and 2nd scores for the pub team members were lower than those achieved by the pundits. For the pub teams and the pundits, the same applies in 2019/20, but this time all three of the pubs' best scores were better than the pundits' best scores. While the bookies' *AE* of 340 was the lowest *AE* score in 2019/20, it would not qualify for the three top player places in either of the pubs and would only make 2nd in the pundits' top scores. While these observations reveal that the bookies and the pundits can be outsmarted by some talented (or lucky) individual member(s) within the crowd, it is not the evidence that would support the central premise of the wisdom of the crowd.

5.2.3 | Differences in AE and CE Over Time

Table 13 lists the *AE* (i.e., $(r - \overline{\rho})^2$) for individual clubs based on predictions from the pub players and pundits. It also includes the bookies' *AE*, but here the difference squared (Diff^2) is the difference in the Bookies' GP and the clubs' final rank, *r*. The promoted clubs are shaded. At the bottom of this table is the percentage contribution of batched ranks, for example, 1–6, 7–12 and so on.

The highlighted cells show those the club error scores that have contributed over 20% of the total *AE* for that group. In 2018/19 the pub players misjudged Wolves: the *GP* for Wolves was 14 and 17 for pubs A and B respectively. Likewise, the pundits did not expect Fulham to be relegated; again the *GP* was much higher at 11th. For 2019/20, it seems that everybody (pub players, pundits, and the bookies) misjudged Sheffield Utd's success while the bookies misjudged Watford and West Ham and Burnley's

TABLE 8 | Group placing, pub A, pub B, pundits, and the bookies, 2018/19 and 2019/20.

	Pub A [20]		Pub B [13]		Pund' [12]		Bookies [23]			Pub A [24]		Pub B [15]		Pund' [9]		Bookies [27]	
2018/19	**GP**	***X* score**	**GP**	***X* score**	**GP**	***X* score**	**GP**	**Diff odds**	**2019/20**	**GP**	***X* score**	**GP**	***X* score**	**GP**	***X* score**	**GP**	**Diff odds**
Man City	1	3	1	2	1	1	1	6	Man City	1	2	1	2	1	2	1	7
Liverpool	2	6	2	4	2	2	2	3	Liverpool	2	4	2	3	2	2	2	9
Man Utd	3	6	3	5	4	6	3	6	Spurs	3	5	3	4	3	1	3	6
Chelsea	5	5	4	4	5	4	4	3	Chelsea	6	7	4	5	4	2	4	8
Spurs	4	7	6	4	3	4	5	5	Man Utd	4	6	6	5	5	3	5	9
Arsenal	6	6	5	4	6	4	6	4	Arsenal	5	6	5	5	6	3	6	7
Wolves	14	13	17	9	10	7	7	4	Wolves	8	9	8	8	9	4	7	8
Everton	7	6	7	6	7	4	8	2	Everton	7	6	7	4	7	3	8	8
Leicester	10	12	10	7	8	6	9	6	Leicester	9	7	10	6	8	4	9	9
West Ham	8	10	9	6	9	7	10	5	West Ham	10	9	9	5	10	5	10	6
Southampton	17	11	14	8	14	8	11	2	Newcastle	16	11	16	9	16	6	11	6
Newcastle	11	10	12	6	15	9	12	2	Aston Villa	15	8	17	6	13	5	12	9
Crystal Palace	13	10	11	6	12	9	13	3	Watford	11	8	11	5	11	2	13	7
Burnley	9	10	8	5	13	7	14	3	Crystal Palace	13	11	12	8	15	5	14	7
Fulham	16	10	15	5	11	5	15	4	Southampton	17	10	15	8	12	4	15	7
Watford	12	10	16	6	18	6	16	3	Bournemouth	12	8	13	7	14	5	16	6
Bournemouth	15	12	13	7	16	6	17	3	Brighton	18	8	18	7	18	5	17	7
Brighton	18	8	18	8	17	6	18	2	Burnley	14	10	14	8	17	8	18	8
Huddersfield	19	8	19	6	19	3	19	4	Sheffield Utd	19	6	20	7	20	3	19	9
Cardiff City	20	8	20	3	20	3	20	5	Norwich	20	6	19	5	19	5	20	9

Note: Shaded clubs are those that were promoted in that year.

TABLE 9 | Pub B, basic scores, and rankings, 2019/2020.

Round 38				Pub (B)—Predict the Prem—2019/2020 season														
Team		**Place**	**Points**	**I**	**H**	**B**	**K**	**O**	**A**	**M**	**D**	**C**	**N**	**G**	**L**	**F**	**J**	**E**	**Group place**
Liverpool	☑	1	99	1	1	1	2	2	2	3	3	2	2	1	1	2	3	1	2
Man City	☑	2	81	2	2	2	1	1	1	1	1	1	1	2	2	1	1	2	1
Man Utd		3	66	4	5	3	6	5	5	7	4	5	5	5	4	5	6	5	6
Chelsea	☑	4	66	5	6	8	5	6	3	5	5	6	3	4	6	3	4	3	4
Leicester		5	62	8	9	13	9	10	11	8	8	9	13	11	8	7	9	10	10
Spurs		6	59	3	3	4	3	4	4	2	2	4	4	3	3	6	2	6	3
Wolves		7	59	6	7	6	8	8	6	10	15	7	8	7	11	17	13	11	8
Arsenal		8	56	7	4	5	4	3	7	4	6	3	6	6	5	4	5	4	5
Sheffield Utd	☒	9	54	16	19	17	20	11	19	14	20	20	18	20	20	20	20	20	20
Burnley		10	54	14	14	9	14	15	14	18	13	15	16	16	17	12	15	14	14
Southampton		11	52	11	12	14	11	17	15	17	16	18	19	14	18	16	11	15	15
Everton		12	49	9	8	7	7	7	8	6	7	8	7	9	7	8	8	9	7
Newcastle		13	44	18	18	19	17	20	18	11	9	16	9	17	9	15	14	17	16
Crystal Palace		14	43	13	13	11	16	18	12	16	12	13	12	10	12	14	19	13	12
Brighton		15	41	17	16	15	18	19	17	20	18	17	17	18	19	11	18	18	18
West Ham	☒	16	39	10	10	12	10	9	9	9	10	10	10	8	10	9	7	8	9
Aston Villa		17	35	19	15	16	15	14	16	19	14	12	15	19	15	18	16	19	17
Bournemouth		18	34	15	17	18	13	13	13	13	17	14	14	13	14	10	10	12	13
Watford	☒	19	34	12	11	10	12	12	10	12	11	11	11	12	13	13	12	7	11
Norwich		20	21	20	20	20	19	16	20	15	19	19	20	15	16	19	17	16	19
On target	☑	*Score*		224	314	336	360	384	386	418	420	422	430	436	436	464	492	514	394
Off course	☒	*Ranking*		1	2	3	4	5	6	7	8	9	10	11	11	13	14	15	n/a
Exact predictions				4	4	6	1	0	1	0	0	1	1	4	2	2	2	3	2

TABLE 10 | Pub B, differences squared, 2019/2020.

Round 38				Pub (B) differences squared (i.e., how your score 'adds up')														
Team		**Place**	**Points**	**I**	**H**	**B**	**K**	**O**	**A**	**M**	**D**	**C**	**N**	**G**	**L**	**F**	**J**	**E**	**Group place**
Liverpool	☑	1	99	0	0	0	1	1	1	4	4	1	1	0	0	1	4	0	1
Man City	☑	2	81	0	0	0	1	1	1	1	1	1	1	0	0	1	1	0	1
Man Utd		3	66	1	4	0	9	4	4	16	1	4	4	4	1	4	9	4	9
Chelsea	☑	4	66	1	4	16	1	4	1	1	1	4	1	0	4	1	0	1	0
Leicester		5	62	9	16	64	16	25	36	9	9	16	64	36	9	4	16	25	25
Spurs		6	59	9	9	4	9	4	4	16	16	4	4	9	9	0	16	0	9
Wolves		7	59	1	0	1	1	1	1	9	64	0	1	0	16	100	36	16	1
Arsenal		8	56	1	16	9	16	25	1	16	4	25	4	4	9	16	9	16	9
Sheffield Utd	☒	9	54	49	100	64	121	4	100	25	121	121	81	121	121	121	121	121	121
Burnley		10	54	16	16	1	16	25	16	64	9	25	36	36	49	4	25	16	16
Southampton		11	52	0	1	9	0	36	16	36	25	49	64	9	49	25	0	16	16
Everton		12	49	9	16	25	25	25	16	36	25	16	25	9	25	16	16	9	25
Newcastle		13	44	25	25	36	16	49	25	4	16	9	16	16	16	4	1	16	9
Crystal Palace		14	43	1	1	9	4	16	4	4	4	1	4	16	4	0	25	1	4
Brighton		15	41	4	1	0	9	16	4	25	9	4	4	9	16	16	9	9	9
West Ham	☒	16	39	36	36	16	36	49	49	49	36	36	36	64	36	49	81	64	49
Aston Villa		17	35	4	4	1	4	9	1	4	9	25	4	4	4	1	1	4	0
Bournemouth		18	34	9	1	0	25	25	25	25	1	16	16	25	16	64	64	36	25
Watford	☒	19	34	49	64	81	49	49	81	49	64	64	64	49	36	36	49	144	64
Norwich		20	21	0	0	0	1	16	0	25	1	1	0	25	16	1	9	16	1
Score				224	314	336	360	384	386	418	420	422	430	436	436	464	492	514	394
Ranking				1	2	3	4	5	6	7	8	9	10	11	11	13	14	15	n/a
Exact predictions				4	4	6	1	0	1	0	0	1	1	4	2	2	2	3	2

Note: Players are denoted 'A' to 'O' and sorted based on their SCORE.

TABLE 11 | Annual results for pub A, pub B, the pundits, and the bookies.

Pub (A)				**Pub (B)**			
2018/19	[20]	2019/20	[24]	2018/19	[13]	2019/20	[15]
1st	126	1st	278	1st	114	1st	224
2nd (2 Players)	142	2nd	334	2nd	128	2nd	314
3rd	162	3rd	336	3rd	154	3rd	336
Highest score	414	Highest score	584	Highest score	394	Highest score	514
% CA	25.8%	% CA	27.9%	% CA	26.9%	% CA	26.7%
DIV	143.7	DIV	89.4	DIV	98.1	DIV	80.0
AE	257.9	AE	416.3	AE	269.2	AE	402.4
CE	114.2	➡CE	326.9	CE	171.1	➡CE	322.4
Pundits						**Bookies**	
2018/19	[12]	2019/20	[9]	2018/19	[27]	2019/20	[23]
Andy Edwards	132	LAST WORD	290	AE ≡ Group score	96	AE ≡ Group score	340
J. Prince-Wright*	168	McNulty [BBC]**	344				
Matt Reed	178	Every Tip	348				
Highest score	386	Highest score	456				
% CA	33.8%	% CA	42.8%				
DIV	95.7	DIV	50.4				
AE	229.5	AE	369.6				
➡CE	133.8	➡CE	319.1				

Note: 2018/19 * The Guardian tied second (168) with J. Prince-Wright. 2019/20 **Phil McNulty tied second (334) with Micheal Owen.

TABLE 12 | Results for the pubs and pundits combined.

2018/19			**2019/20**		
Place	**Score**	**Team**	**Place**	**Score**	**Team**
1st	114	Pub B	1st	224	Pub B
2nd	126	Pub A	2nd	278	Pub A
3rd	128	Pub B	3rd	290	Pundits
4th	132	Pundits	4th	314	Pub B

	CE	**AE**	**DIV**	***n***
2018/19	122.6	253.6	131.0	45
2019/20	321.1	403.2	82.1	48

Error per club

2018/19			**2019/20**		
Place	**EPC**	**Team**	**Place**	**EPC**	**Team**
1st	2.39	Pub B	1st	3.35	Pub B
2nd	2.51	Pub A	2nd	3.73	Pub A
3rd	2.53	Pub B	3rd	3.81	Pundits
4th	2.57	Pundits	4th	3.96	Pub B

EPC	**CE**	**AE**
2018/19	2.48	3.56
2019/20	4.01	4.49

performances, albeit to a smaller extent. Table 12 reveals an absence of higher difference scores in 2018/19 for both pub A and the pundits, suggesting that there were fewer upsets in 2018/19 when compared with 2019/20.

The batched contributions show that nearly all players were justified in predicting the 'Big Six' clubs would dominate the top rank positions, and thus their contribution to the *AE* is by far the lowest. This suggests that the public's perception of the EPL is a league that is not competitively balanced. The highest contribution to *AE* is from the 7 to 12 batch, which reflects the spread of predictions in the prediction matrices (Table 5).

The evidence presented here indicates that the volatility in Page's (2007a) DPT calculations suggest that predicting unexpected outcomes is not a strong feature of the wisdom of the crowd, which raises questions on the concept's broader usefulness. Moreover, as emphasised by Lorenz et al. (2011, 9020), the wisdom of the crowd effect is based on an aggregation of individual estimates and therefore it is a statistical phenomenon rather than a pure social psychological effect.[6]

6 | Conclusions

This paper outlined a game that asked players to predict the end of season rank positions of clubs in the EPL. Scoring in the game shares similarities with Page's (2007a) DPT and it offered us the

TABLE 13 | Difference scores for all groups, 2018/18 and 2019/20.

	Plc	Pub A [20]	Pub B [13]	Pund′ [12]	Bookies [23]			Plc	Pub A [24]	Pub B [15]	Pund′ [9]	Bookies [27]	
2018/19	r	$(r-\overline{\rho})^2$	$(r-\overline{\rho})^2$	$(r-\overline{\rho})^2$	**GP**	**Diff 2**	**2019/20**	r	$(r-\overline{\rho})^2$	$(r-\overline{\rho})^2$	$(r-\overline{\rho})^2$	**GP**	**Diff 2**
Man City	1	0.6	0.3	0.0	1	0	Liverpool	1	1.0	1.2	0.8	2	1
Liverpool	2	2.0	0.7	0.2	2	0	Man City	2	0.6	0.6	0.8	1	1
Chelsea	3	3.6	2.5	3.8	4	1	Man Utd	3	4.9	4.6	3.9	5	4
Spurs	4	2.6	2.8	1.2	5	1	Chelsea	4	7.3	2.7	1.3	4	0
Arsenal	5	1.6	1.3	0.9	6	1	Leicester	5	24.3	23.6	9.3	9	16
Man Utd	6	8.4	5.5	4.4	3	9	Spurs	6	7.6	7.5	9.0	3	9
Wolves	7	61.7	92.2	32.9	7	0	Wolves	7	15.1	16.5	4.6	7	0
Everton	8	6.2	2.9	2.1	8	0	Arsenal	8	11.7	11.4	7.0	6	4
Leicester	9	16.3	9.8	3.5	9	0	*Sheffield Utd*	9	91.8	92.7	110.0	19	100
West Ham	10	12.0	6.6	13.3	10	0	Burnley	10	24.9	23.6	34.7	18	64
Watford	11	13.7	22.6	37.3	16	25	Southampton	11	27.9	22.3	9.4	15	16
Crystal Palace	12	13.3	3.8	13.8	13	1	Everton	12	27.3	19.5	16.7	8	16
Newcastle	13	13.7	6.1	12.3	12	1	Newcastle	13	13.3	18.3	18.2	11	4
Bournemouth	14	10.3	4.0	5.6	17	9	Crystal Palace	14	7.3	6.3	7.0	14	0
Burnley	15	21.5	37.1	11.6	14	1	Brighton	15	10.8	9.0	8.9	17	4
Southampton	16	11.5	17.9	12.5	11	25	West Ham	16	37.6	44.9	33.7	10	36
Brighton	17	7.0	8.9	6.3	18	1	*Aston Villa*	17	9.4	5.3	9.4	12	25
Cardiff City	18	5.5	2.6	3.4	20	4	Bournemouth	18	28.0	23.2	20.1	16	4
Fulham	19	30.9	24.8	61.0	15	16	Watford	19	60.7	61.9	60.7	13	36
Huddersfield	20	15.8	16.6	3.6	19	1	*Norwich*	20	4.8	7.4	4.1	20	0
	Total AE	257.9	269.2	229.5	Total	96		Total AE	416.3	402.4	369.6	Total	340
Contribution to AE							Contribution to AE						
	1–6	7.2%	4.9%	4.5%	1–6	12.5%		1–6	17.7%	14.9%	10.9%	1–6	9.1%
	7–12	47.8%	51.3%	44.8%	7–12	27.1%		7–12	77.0%	69.1%	79.4%	7–12	58.8%
	13–17	24.8%	27.5%	21.1%	13–17	38.5%		13–17	30.4%	31.1%	33.6%	13–17	20.3%
	18–20	20.2%	16.3%	29.6%	18–20	21.9%		18–20	36.3%	34.3%	37.0%	18–20	11.8%

opportunity to apply and evaluate Page's theorem using a natural experiment. In the process, we created prediction matrices that helped to understand the factors driving DIV used in the DPT, and we created two ratios of consensus as corollaries to the DPT's diversity score to help understand the meaning of diversity in the context of Page's (2007a) DPT. The game looked at the predictions of two pub-based teams and compared their predictions to those made by nationally recognised football commentators (pundits) and the league rankings inferred by the odds offered by national bookmakers (bookies). For the purpose of this experiment, the pundits and the bookies were seen as experts. Our results shine a light on three aspects of this experiment: first, from the perspective of the DPT; second, in terms of the performance of the players; and third, the differences in DPT scores across two football seasons.

In the context of the DPT, our results reveal that at a practical level the application of DPT is more amenable to estimates of single issues predictions (e.g., the number of beans in a jar) and less amenable to estimates involving multiple interdependent predictions because the crowd's error cannot be interpreted to provide the best answer. This outcome means that practical applications of the DPT are limited. Our results also reveal that in all but one instance there were experts who performed better than the crowd such that we cannot conclude that the wisdom of the crowd has prevailed or that 'diversity will almost always trump ability' (Page 2007b, 11). Moreover, in terms of the predictions between the groups at an aggregate level and based on the AE, experts' predictions were closer to the final outcome than those predicted by groups of pub players, which is a result that is counter to the predictions of the DPT or the wisdom of the crowd more generally. Even here however there were individual pub players whose predictions were better than the pundits and the bookies, which again does not support the notion that the wisdom of the crowd as an all-encompassing and durable phenomenon.

We showed that the predictions of all players overlooked the possibility of surprising results. At a game level this revealed itself in higher AE and crowd error scores where particular football teams did better or worse than expected. This outcome is a reflection of the mindsets of our game players and less on the efficacy of the wisdom of the crowd as a concept, but it raises questions relating to how people might view uncertainty or unexpected events. Finally, the arguments for diversity are laudable if not compelling, but the algebraic proof for the role of diversity as offered by Page's (2007a) DPT is at best limited to specific single event outcomes and could be a distraction or worse misunderstood. Page's (2007a) contribution lies in his recommendations on how to manage and nurture diversity, which are worthy on their own merit. Future research could examine further instances where the wisdom of the crowd may be strong, such as predicting economic growth or inflation rates.

Endnotes

[1] Note that the city of Bristol has never had an EPL football team. Although no set of pub customers are likely to ever be perfectly independent of each other, our set of pub customers may have a lower level of connection with EPL football teams than as anywhere else in England simply because of their city being relatively large but still not have a EPL football team.

[2] In practice individual players submitted their predictions in early September of each football season which was after 4 or 5 games of a 38-game season.

[3] In both years there were no equal ranks that had to be manually adjusted. In addition, adding up the percentages by column gives an indication of the bookies mark-up, in 2018/19 the average was 18.6%, in 2019/20 it was 10.8%.

[4] Thurstonian modelling describes the mapping of a continuous scale onto a discrete possibly ordered categorical response. Although that modelling approach accounts for non-independence of alternatives, it would compound the complexity in making predictions due to the effects of social interactions (felt or otherwise) between players, which we think are likely to be important directly and indirectly when making predictions. Therefore, we simplify the issue and examine a set of ratios to test the generalisability of the wisdom of the crowd premise.

[5] This is the same outcome when the pub teams and pundits are combined as one crowd in each season.

[6] There is expected to be a social effect present in our results as our pundits, bookies, and probably even our players discussed football related to these issues in a social setting, and this may have influenced their perceptions and predictions, thereby rendering them not statistically independent. This reflects reality in many circumstances where the wisdom of the crowd plays out in real life.

References

Galton, F. 1907. "Vox Populi." *Nature* 75: 450–451.

Granger, C. W. J., and M. H. Pesaran. 2000. "Economic and Statistical Measures of Forecast Accuracy." *Journal of Forecasting* 19, no. 7: 537–550.

Lorenz, J., H. Rauhut, F. Schweitzer, and D. Helbing. 2011. "How Social Influence Can Undermine the Wisdom of the Crowd." *Proceedings of the National Academy of Sciences of the United States of America* 108, no. 22: 9020–9025.

Madden, P., and T. Robinson. 2012. "Supporter Influence on Club Governance in a Sports League: A Utility Maximisation Model." *Scottish Journal of Political Economy* 59, no. 4: 339–360.

Nobre, D. A., and J. F. Fontanari. 2020. *Prediction Diversity and Selective Attention in the Wisdom of Crowds*. Cornell University. https://arxiv.org/abs/2001.10039.

Page, S. E. 2007a. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton University Press.

Page, S. E. 2007b. "Making the Difference: Applying a Logic of Diversity." *Academy of Management Perspectives* 21, no. 4: 6–20.

Prinz, A., and S. Thiem. 2021. "Value-Maximising Football Clubs." *Scottish Journal of Political Economy* 68, no. 5: 605–622.

Reade, J. J., C. Singleton, and A. Brown. 2021. "Evaluating Strange Forecasts: The Curious Case of Football Match Scorelines." *Scottish Journal of Political Economy* 68, no. 2: 261–285.

Sloane, P. J. 1971. "The Economics of Professional Football: The Football Club as a Utility Maximiser." *Scottish Journal of Political Economy* 18, no. 2: 121–146.

Surowiecki, J. 2004. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. Doubleday.

Triguero-Ruiz, F., and A. Avila-Cano. 2023. "On Competitive Balance in the Group Stage of the UEFA Champions League." *Scottish Journal of Political Economy* 70, no. 3: 231–248.

Yudkowsky, E. 2007. "The Error of the Crowds." Blog. Accessed September 09, 2020. https://www.lesswrong.com/posts/zAvhTnQX6ynJF7pyh/the-error-of-crowds.

Appendix A

Table of Definitions and Equations

Page (2007a) Diversity Prediction Theorem

Crowd error (*CE*) is $\sum(r-\overline{p})^2$ where (r) is the ranked place of a particular football club at the end of a season (e.g., 6th) and ($\overline{p}$) is the average of all the player's place predictions (e.g., 3.71) for that club, thus the difference between (r) and ($\overline{p}$) is 2.29. (see Table 1).

Average Error (*AE*) is the sum of the average of the squares of the difference between each club's actual ranked place (r) and each player's predicted place (p): $\sum (r-p)^2$.

Diversity (*DIV*) is the population variance (σ^2) in predictions (p): $\sigma^2 = \sum (x_i - \overline{x})^2/n$.

Total Crowd Error (Σ*CE*), Average Error (Σ*AE*), and Diversity (Σ*DIV*) are the sum of each parameter per club, which become the basis of Page (2007a) theorem: $CE \equiv AE - DIV$.

Error on a per-rank basis (*EPR*) is the square root of the value of either the average Crowd Error $\sqrt{(\Sigma CE/n)}$ or the average of the Average Error $\sqrt{(\Sigma AE/n)}$, where n = number of people making the predictions. In games predicting the end-of-season league places of football clubs, and when applied to a player's total score (see below), the *EPR* offers a more intuitive interpretation of how close the player's place predictions are to the actual outcomes.

Our Game

Players Total Score (*TS*)

In our game the player's total score is the same as the Total Average Error (see above and Table 10), with a 20-club league it would be:

$$TS = \sum_{i=1}^{20} (p_i - r_i)^2$$

In a game of predicting league places, the player with the lowest *TS* (and lowest *EPR*) would be the winner.

Group Place (*GP*)

In the context of predicting the end-of -season places in a sporting league, *GP* is the rank of the average of the players predictions for any given club ($\overline{p}$). As an easily understood (albeit crude) integer the *GP* allows players to roughly evaluate their individual rank predictions relative to those made by the group.

Understanding Diversity (or Variation) in Predictions

Prediction matrices show the distribution of actual (or unique) predictions relative to all possible (or potential) predictions. The area of the matrix is defined by the number of clubs (***c***) in a league and the number of places (r) in the league, for example 20 clubs × 20 league places: ($c \times r$).

These matrices can be read by column (league rank 1 to 20) or by row (football club rank position). The sum of the rows and the sum of the columns are always equal and add up to the number of players in a group (see *X* and *Z* scores below).

Potential predictions (PP) are the total number of potential predictions for each club in each league, i.e., the square of the number of clubs (c): $PP = c^2$.

Unique predictions (UP) are the total number of predictions made by a group of players and are determined by the number of players making predictions (pp) and the number of clubs (c):

$$UP = pp \times c$$

It follows that *UP* can be greater than *PP* when the number of predicting players is greater than the number of clubs in the league (and vice versa). See Table 7 for example values of *UP* and *PP*.

***X*-score per club** is the number of different places or unique place predictions per club. Thus, if a group of players ($n = 15$) predict that a particular club will finish either 1st or 2nd that club's *X*-score would be 2. The *X*-score is a simple way to understand the variation (*DIV*) in predictions per club (compare Man City and Crystal Place, Table 5).

***Z*-scores per league place** is the number of different club predictions per league (rank) position (e.g., 1st, 2nd, 3rd and so on). Thus, if a group of players predicts that there are only two clubs that might end up in 1st place, that place's *Z*-score would be 2. The *Z*-score is a simple way to infer the variation in predictions for a particular league place (see 1st place, Table 5).

Σ(*X*-scores) and Σ(*Z*-scores) the sum of the *X* and *Z* scores will always be equal, and they show where the players predictions overlap: the lower the sum of the *X* and *Z* scores, the greater the overlap in predictions (or consensus) on the club's league place (see below).

Complete agreement (*CA*) In the prediction matrices (Tables 4 and 5), the cell values along the diagonal from coordinates 1:1 to 20:20 indicate the degree of complete agreement (*CA*). For any club, *CA* is the number of players whose predictions agree with the group place (*GP*) (see above). The summation of all these club specific *CA* values (Σ*CA*) gives the total number of times that the players predictions agreed with the group place (*GP*).

Ratio #1 is the ratio of unique (actual) predictions (*UP*) to potential predictions (*PP*), expressed as a percentage: $(UP/PP) \times 100$. The higher this percentage, the greater the overlap in predictions (or consensus) on the club's league place (see below).

Ratio #2 is the ratio of the sum of predictions along the line of complete agreement (Σ*CA*) to the sum of all predictions $(\Sigma CA/PP) \times 100$. The higher this percentage, the greater the agreement with the group place (*GP*).