TI 2018-101/IV
Tinbergen Institute Discussion Paper

Incentives, Performance and Choking in Darts

Bouke Klein Teeselink[1]
Rogier J. D. Potter van Loon[2]
Martijn (M.J.) van den Assem[1]
Dennie van Dolder[1]

[1] VU Amsterdam, The Netherlands
[2] Erasmus University, The Netherlands

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

Incentives, Performance and Choking in Darts

Bouke Klein Teeselink[1,2], Rogier J.D. Potter van Loon[3], Martijn J. van den Assem*[1,2] and Dennie van Dolder[1,2,4]

[1] *School of Business and Economics, VU Amsterdam*
[2] *Tinbergen Institute*
[3] *Erasmus School of Economics, Erasmus University Rotterdam*
[4] *Centre for Decision Research and Experimental Economics, University of Nottingham*

December 19, 2018

Abstract

This paper examines the effect of incentives on the performance of darts players. We analyze four data sets comprising a total of 123,402 darts matches of professional, amateur, and youth players. The game of darts offers an attractive natural research setting, because performance can be observed at the individual level and without the obscuring effects of risk considerations and the behavior of others. We find that amateur and youth players perform better under moderately higher incentives, but choke when the incentives are really high. Professional players similarly display better performance under higher incentives, but appear less susceptible of choking. These results speak to a growing literature on the limits of increasing incentives as a recipe for better performance.

Keywords: incentives; choking under pressure; performance; darts
JEL Classifications: D01; D91; Z20

*Corresponding author. VU Amsterdam, School of Business and Economics, De Boelelaan 1105, 1081 HV Amsterdam, the Netherlands. E-mail: m.j.vanden.assem@vu.nl. The authors thank Roger Boyesen of Darts for Windows for providing and explaining the data. The paper has benefited from discussions with participants of FUR 2018 York, SABE/IAREP 2018 London, TIBER 2018 Tilburg, RBFC 2018 Amsterdam, and JDMx 2018 Konstanz.

1 Introduction

Incentives are at the core of economics. From labor supply to crime, and from consumption to education, incentives play a central role in economic theory. An important prediction is that people exert more effort when they face stronger incentives, and that this increased effort in turn leads to better performance. Various experimental studies show that higher monetary incentives indeed improve performance (Smith and Walker, 1993; Camerer and Hogarth, 1999; Gneezy and Rustichini, 2000). Outside the behavioral laboratory, research has similarly shown, for example, that workers produce more output when they are paid a piece rate rather than an hourly rate (Lazear, 2000; Shearer, 2004; Paarsch and Shearer, 1999) and that students perform better at tests when they are paid according to their performance (Levitt et al., 2016).

Psychological research, however, suggests that higher incentives do not always improve performance. First, they can cause people to consciously think about their actions in otherwise automatically performed tasks, and thereby impede performance (Langer and Imber, 1979; Baumeister, 1984; Camerer et al., 2005). Second, higher incentives can lead to more arousal. The Yerkes-Dodson law postulates that arousal has a non-monotonic effect: moderate increases enhance performance, whereas the effect turns negative if arousal surpasses a critical threshold that depends on task difficulty (Yerkes and Dodson, 1908). Ariely et al. (2009) show that high stakes harm performance in a wide range of experimental tasks. Inferior performance in the presence of incentives for superior performance is known as choking under pressure (Baumeister and Showers, 1986; Beilock, 2010).

The present paper contributes to the understanding of the link between incentives and performance by analyzing four large data sets of competitive darts matches of professional, amateur and youth players. Darts offers a unique combination of attractive features for this type of research. The game is played in a real-world natural environment. Like many other real life activities—such as the work of pilots, bus drivers, soldiers, surgeons and dentists—the task is neither entirely physical nor entirely mental, but combines elements of both. Just as with controlled experiments, the game has a clearly defined set of rules and objectives. Performance can be observed at the individual level, and unlike many other field tasks it is not confounded by the behavior of others, such as colleagues and competitors. As opposed to many other field settings—where it can be rational for risk-averse individuals to shift to low-reward low-risk strategies when the stakes increase—the optimal approach to darts is independent of risk preferences.

By using darts data, we connect to a broader literature that uses sports data to investigate economic hypotheses. Examples are the study of discrimination in basketball (Price and Wolfers, 2010), principal-agent theory in cricket (Gauriot and Page, 2015), and mixed-strategy Nash equilibria in soccer penalty shootouts (Chiappori et al., 2002). Choking under pressure has previously been observed in, for example, penalty shootouts in soccer (Dohmen, 2008), free throws in basketball (Cao et al., 2011; Toma, 2017), putting in golf (Hickman and Metz, 2015), shooting in biathlon (Harb-Wu and Krumer, 2017; Lindner, 2017), and in tennis (Paserman, 2010; Cohen-Zada et al., 2017).

To investigate the effect of incentives on performance in darts, we exploit naturally occurring within-match variation in the benefit (cost) of throwing well (poorly). We find that amateur and youth players perform better under moderately higher incentives, but choke when the incentives are really high. Professional players similarly display better performance under higher incentives, but appear less susceptible of choking.

In the remainder of the paper we explain the game of darts and our data (Section 2), present the analyses and results (Section 3), and conclude (Section 4).

2 Description Darts and Data

2.1 Darts

In darts, two players compete with each other by sequentially throwing darts at a dartboard. A dartboard is divided into areas that represent points in the range of 1 to 20 (see Figure 1). When a dart is thrown in the outer or inner narrow band, the number of points is doubled (outer) or tripled (inner). The outer ring in the center of the dartboard (outer bull) gives 25 points, the inner ring (bullseye) gives 50 points. The maximum score with one dart equals 60 points (triple 20).

Darts matches are played in either 'leg' or 'set' format. A match in leg format is a best-of-n contest, where each of the n sub-contests is called a leg. A match in set format also is a best-of-n contest, but there each sub-contest is called a set, which

Figure 1: Dartboard

in turn is a best-of-n contest with legs. A set is thus identical to a match in leg format.

Players commonly start a leg with 501 points each, and take turns to throw three darts. One turn of three darts is generally referred to as one ‘throw’. The sum of the points in a throw is subtracted from the remaining number of points. To finish and win a leg, a player is required to reach zero by hitting either a ‘double’ or the bullseye. For example, a player with 18 points remaining can finish by throwing double 9. If the score of a dart exceeds the number of points the player has left, her entire throw of three darts is rendered invalid. Players take turns to start legs. The starter of the first leg is generally determined by shots at the bullseye, with the player closest starting.

2.2 Data

Our data are from Darts for Windows (www.dartsforwindows.com). Darts for Windows collects data from various sources, most notably from darts associations and darts competitions that use the Darts for Windows computer software. There are four categories of data: Youth, Super League, British Inter-County Championship

(BICC), and International. We downloaded all available BICC data on November 26, 2017, and all available data for the other three categories on July 11, 2017.

The four data sets cover the matches of different types of players, ranging from amateur youth players to professional adults. The youth sample consists of 1,671 matches from tournaments for boys under 18, boys under 21 and girls under 21, that took place in the period 2001-2017. The Super League is a regional amateur league that is played mostly in the United Kingdom. Our sample contains 1,516 matches from 2007-2017. The BICC is a competition between amateur players from various counties in the United Kingdom. The BICC sample comprises 10,369 matches played in the period 2005-2017. The international tournaments sample covers 15,205 matches between professional players from 1974-2017, and includes matches played at famous tournaments such as the UK Open and the PDC World Darts Championship.

For each match, we have granular data down to each player's score in one throw of three darts. Along with the score per throw, we know the date of the match, players' names, and the starter of the first leg.

In our analyses we treat sets as separate matches. Most matches in the Youth (51%), Super League (92%), and BICC (100%) sample are played in set format. These matches are virtually always played between two teams, where each set is played by a different team member and sets can rightfully be regarded as matches on their own. In the International sample, only a small proportion (5.6%) of the matches are in set format, and sets are generally played by the same player. For consistency, we nevertheless similarly treat these sets as separate matches. To make sure that our results are not sensitive to this approach, we also conduct robustness analyses that exclude the data from international matches that were played in set format. Treating sets as matches increases the total number of matches in the four samples combined from 28,761 to 124,072.

We exclude all matches where legs do not start at 501 points, matches where one or more scores are missing, matches with more than two players, and matches where both players have the same name. After these cleaning operations, 123,402 matches remain.

Table 1 presents summary statistics. In total, our data comprises more than eight million throws. There are clear skill differences across the four categories, with average points per throw ranging from 49 in youth tournaments to 66 in international

Table 1: Summary statistics

	Youth	Super League	BICC	International
Matches	$5,216$	$11,309$	$89,019$	$17,855$
Legs	$14,869$	$47,020$	$382,195$	$85,446$
Throws	$279,295$	$746,383$	$5,922,129$	$1,200,383$
Players	$2,551$	$4,675$	$10,788$	$2,644$
Legs per match	2.85	4.16	4.29	4.79
Throws per leg	18.78	15.87	15.50	14.05
Points per throw (all)	49.26	58.13	60.06	65.84
Points per throw (first three)	59.31	67.70	70.58	79.62
Starters winning the leg (proportion)	0.55	0.56	0.57	0.59
One-dart finish opportunities	$33,454$	$73,191$	$587,876$	$111,241$
Successful one-dart finishes (proportion)	0.32	0.43	0.42	0.47

Notes: Matches, *Legs*,*Throws* and *Players* are the number of matches, legs, throws, and players, respectively. *Legs per match* is the average number of legs per match. *Throws per leg* is the average number of throws by both players combined per leg. *Points per throw (all)* is the average number of points per throw across all throws. *Points per throw (first three)* is the average number of points per throw across players' first three throws in every leg. *Starters winning the leg* is the proportion of legs won by the player who started the leg. *One-dart finish opportunities* is the number of throws where a player can finish the leg with one dart. *Successful one-dart finishes* is the proportion of throws where a player could finish the leg with one dart and finished the leg with one, two or three darts.

tournaments. These skill differences are also reflected in the average leg length, with better players taking fewer throws to finish a leg.

3 Analyses and Results

The incentive to do well in darts varies both across and within matches. Across matches, players will be motivated by the amount of prestige, prize money, and media attention. Compare, for example, the final of the internationally televised PDC World Darts Championship where the 2018 winner took home £400,000, with a match in the first round of the men's singles tournament of the Lincolnshire Family Darts Festival where the 2018 winner cashed £1,000. Our data, however, does not provide sufficient information to systematically proxy for such variation.

Within matches, the incentive to throw well increases moderately when both players are close to winning a leg, and considerably when both players are close to winning the match. Our analyses exploit this naturally-occurring within-match variation, and consider its effect on players' finishing performance (Section 3.1) and on the points they throw in the first three throws of a leg (Section 3.2).

3.1 Finishing

We first examine how players' finishing performance is affected when their opponent can finish in the subsequent throw with one, two, or three darts. This required number of darts inversely proxies for the probability that the opponent will finish, and thus for the incentive of the throwing player to do well. The pressure is relatively low when the opponent cannot finish, and increases with the ease with which the opponent can finish. Second, we examine how finishing performance is affected if both players can win the match by winning the current leg. Such legs are consequential because poor performance can irreversibly result in losing the match, whereas strong performance can secure the win.[1] Last, we consider the interaction of situations where both can win the match by winning the current leg and situations where the opponent can finish in the subsequent throw. In such critical cases, the pressure to do well can be assumed to be extraordinarily high.

We exclusively consider finishing performance when the player can finish the leg with one dart. This is the case when she has 2, 4, ..., 38, 40 or 50 points left. Strategy plays no role in these situations, because any approach other than trying to finish the leg in the current throw is sub-optimal. In contrast, if a player needs multiple darts to finish she might instead try to maneuver herself into a better finishing position for the next throw. Such a strategy can be attractive in situations where the opponent is unlikely or unable to finish, and would thus lead to the false impression of lower performance in lower-incentive situations.

We use a fixed-effects logit model to regress finishing performance on incentives. We control for skill differences between players and through time by including player-match fixed effects, and for possible warming-up and fatigue effects within a match by including a polynomial of order n for the player's number of throws in the match prior to the current throw, where the value of n is chosen to minimize the AIC.[2]

Table 2 presents the regression results. Model 1 shows how a player's finishing

[1] We intentionally only consider situations where *both* can win the match, not those where only one can win. First, if only one can win, the incentive can be both high and low, depending on the closeness of the match. Second, using situations where only one can win would lead to biased coefficient estimates due to regression to the mean. A player who is ahead (behind) because of exceptionally strong (weak) performance is likely to subsequently do worse (better), which confounds the causal effect of being ahead (behind) on performance.

[2] If the match has a set format, the number of throws includes the throws in previous sets played by the same player.

Table 2: Regression results for the likelihood of finishing

	Model 1	Model 2	Model 3
Panel A: International			
Opp. can finish with 1 dart	0.114***		0.115***
	(0.000)		(0.000)
Opp. can finish with 2 darts	0.073***		0.073***
	(0.000)		(0.000)
Opp. can finish with 3 darts	0.028***		0.028***
	(0.002)		(0.002)
Both can win match		−0.031**	−0.007
		(0.014)	(0.872)
Opp. can finish with 1 dart × Both can win match			−0.034
			(0.431)
Opp. can finish with 2 darts × Both can win match			−0.017
			(0.698)
Opp. can finish with 3 darts × Both can win match			0.003
			(0.943)
Player-match fixed effects	Yes	Yes	Yes
Polynomial throw number (order)	3	3	3
McFadden pseudo R-squared	0.356	0.353	0.356
Observations	111,241	111,241	111,241
Panel B: BICC			
Opp. can finish with 1 dart	0.120***		0.129***
	(0.000		(0.000)
Opp. can finish with 2 darts	0.075***		0.080***
	(0.000		(0.000)
Opp. can finish with 3 darts	0.037***		0.041***
	(0.000		(0.000)
Both can win match		−0.100***	−0.006
		(0.000	(0.642)
Opp. can finish with 1 dart × Both can win match			−0.125***
			(0.000)
Opp. can finish with 2 darts × Both can win match			−0.073***
			(0.000)
Opp. can finish with 3 darts × Both can win match			−0.055***
			(0.000)
Player-match fixed effects	Yes	Yes	Yes
Polynomial throw number (order)	5	5	5
McFadden pseudo R-squared	0.324	0.322	0.325
Observations	587,876	587,876	587,876

Notes: The table reports the average marginal effects resulting from logit regression analyses of finishing performance across all throws where the player can finish the leg with one dart. The dependent variable takes a value of 1 if the player finishes in the given throw, and 0 otherwise. *Opp. can finish with 1 dart* (*2 darts*; *3 darts*) is a dummy variable that takes the value of 1 if the player's opponent can finish the leg in the subsequent throw with one dart (two darts; three darts). *Both can win match* is a dummy variable that takes a value of 1 if both players can win the match by winning the current leg. The regression specifications include player-match fixed effects, and a polynomial of order n for the player's prior number of throws, where n is chosen to minimize the AIC. Average marginal effects are corrected for incidental parameter bias (Hahn and Newey, 2004). p-values are in parentheses. Asterisks denote significance at the 0.01 (***), 0.05 (**) and 0.1 (*) level.

Table 2: Continued

	Model 1	Model 2	Model 3
Panel C: Super League			
Opp. can finish with 1 dart	0.124***		0.131***
	(0.000)		(0.000)
Opp. can finish with 2 darts	0.077***		0.079***
	(0.000)		(0.000)
Opp. can finish with 3 darts	0.037***		0.040***
	(0.001)		(0.000)
Both can win match		−0.045***	0.028
		(0.001)	(0.458)
Opp. can finish with 1 dart × Both can win match			−0.103**
			(0.011)
Opp. can finish with 2 darts × Both can win match			−0.048
			(0.240)
Opp. can finish with 3 darts × Both can win match			−0.048
			(0.303)
Player-match fixed effects	Yes	Yes	Yes
Polynomial throw number (order)	5	5	5
McFadden pseudo R-squared	0.316	0.313	0.316
Observations	73,191	73,191	73,191
Panel D: Youth			
Opp. can finish with 1 dart	0.145***		0.159***
	(0.000)		(0.000)
Opp. can finish with 2 darts	0.066***		0.070***
	(0.000)		(0.000)
Opp. can finish with 3 darts	0.032**		0.031*
	(0.043)		(0.064)
Both can win match		−0.087***	−0.020
		(0.000)	(0.645)
Opp. can finish with 1 dart × Both can win match			−0.109**
			(0.015)
Opp. can finish with 2 darts × Both can win match			−0.038
			(0.408)
Opp. can finish with 3 darts × Both can win match			0.009
			(0.856)
Player-match fixed effects	Yes	Yes	Yes
Polynomial throw number (order)	5	6	5
McFadden pseudo R-squared	0.332	0.327	0.333
Observations	33,454	33,454	33,454

probability changes if her opponent can finish the leg with one, two or three darts in the next throw, compared to situations where her opponent cannot finish. In all four data sets, higher incentives consistently lead to better performance. A player's finishing probability increases by 3-4 percentage points if her opponent can finish with three darts, by 7-8 percentage points if her opponent can finish with two darts, and by 11-14 percentage points if her opponent can finish with only one dart.

Model 2 shows that a player's finishing probability on average tends to deteriorate by 3-10 percentage points if both players can win the match by winning the leg, compared to situations where none or only one player is close to winning the match. This negative effect is significant at the one-percent level for BICC, Super League and Youth matches, and at five percent for International matches. For the interpretation of this result it is important to note that in many of these cases, the incentive to perform well was particularly high, because often in addition the opponent was able to finish the match with her next throw (93-94% in all four samples).

Model 3 disentangles the marginal effects of the opponent's ability to finish the leg, legs that give both players the opportunity to win the match, and the interaction of these two. When both types of incentive variables and their interaction are included, the marginal effects of the opponent's ability to finish still closely resemble those of Model 1: players improve their finishing performance by 3-16 percentage points if their opponent is able to finish. Compared to Model 2, the marginal effects of both players being able to win the match are no longer significantly negative. This is not very surprising, because these marginal effects now refer to situations where the player is significantly ahead in the leg: the player can finish with one dart, whereas the opponent is unable to finish in the next throw. In such situations, players do not have a particularly strong incentive to perform well.

When the incentive to perform well is really high—the opponent is able to finish the leg in the next throw with just one dart and at the same time the leg gives both players the opportunity to win the match—BICC, Super League, and Youth players show evidence of choking. When the opponent can finish with one dart, players' finishing performance is significantly worse in decisive legs where both players can win the match than in non-decisive legs. More specifically, in decisive legs they are 13.1 (BICC), 7.5 (Super League) and 12.9 (Youth) percentage points less likely to hit the double.[3] In the category of international matches, where players are more skilled and more experienced, there is no compelling evidence of choking in such situations.

One possible concern about the previous results is that a player's finishing difficulty likely correlates with the opponent's opportunity to finish in the next throw. Players tend to have fewer points left when their opponent can finish with fewer darts. Because finishing difficulty decreases with the number of points left, players can be expected to have a lower chance of finishing if their opponent can finish with

[3] BICC: $0.6 + 12.5$; Wald $\chi^2(1) = 488.04$, $p < 0.001$. Super League: $-2.8 + 10.3$; Wald $\chi^2(1) = 18.17$, $p < 0.001$. Youth: $2.0 + 10.9$; Wald $\chi^2(1) = 40.15$, $p < 0.001$.

fewer darts.[4] Any such correlation would thus work counter to our findings. In the regressions we can nevertheless control for finishing difficulty by adding fixed effects for the number of points that the player has left at the start of her throw. A downside of this approach is that the player's number of points left correlates with the opponent's number of points left, and the points-left fixed effects can therefore be expected to capture part of the incentive effect of the opponent's ability to finish with one, two, or three darts. Table A1 in Appendix A gives the new results, and shows that these are similar to what we found before. The only material difference is that the marginal effects of the opponent's ability to finish the leg are systematically smaller, which likely reflects the above-mentioned downside of adding points-left fixed effects.

Another possible concern relates to our treatment of set-format matches and the absence of evidence for choking in international matches. Throughout our analyses, we have treated sets as separate matches, which especially makes sense if each set is played by a different player of a team. In international tournaments, however, set-format matches are generally entirely played between the same two players, and winning a set is therefore substantially less important than winning the match. Treating sets as matches may consequently have diluted possible evidence of choking effects in this category of data. Table B1 in Appendix B shows the results for the International sample after excluding all set-format matches. The results are virtually unchanged.

3.2 Points Thrown

We now turn to the early stage of legs, to examine how points-throwing performance is affected if both players can win the match by winning the current leg. As explained in the previous section, such legs are consequential because poor performance can irreversibly result in losing the match, whereas strong performance can secure the win. We consider each player's first three throws only. Strategy plays no role in these first few throws, because any approach other than trying to throw as many points

[4] Compare, for example, a player with 2 and a player with 40 points remaining. Both can finish with one dart (double 1 and double 20, respectively). If the player with 2 points misses and instead throws 1 point, her throw is over. If the player with 40 points misses and instead throws 20 points, she still has the opportunity to finish with double 10.

Table 3: Regression results for the number of points thrown

	Model 1: Throw 1	Model 2: Throw 2	Model 3: Throw 3	Model 4: Throws 1-3
Panel A: International				
Both can win match	-0.246	-0.992^{**}	-0.432	-0.562^{**}
	(0.612)	(0.042)	(0.373)	(0.045)
Player-match fixed effects	Yes	Yes	Yes	Yes
Polynomial throw number (order)	9	5	9	7
Observations	170,892	170,892	170,884	512,668
Adjusted R-squared	0.234	0.238	0.232	0.241
Panel B: BICC				
Both can win match	-1.732^{***}	-1.943^{***}	-1.922^{***}	-1.791^{***}
	(0.000)	(0.000)	(0.000)	(0.000)
Player-match fixed effects	Yes	Yes	Yes	Yes
Polynomial throw number (order)	9	9	9	9
Observations	764,390	764,390	764,384	2,293,164
Adjusted R-squared	0.129	0.130	0.131	0.134
Panel C: Super League				
Both can win match	-1.646^{***}	-0.703	-0.352	-0.992^{***}
	(0.004)	(0.212)	(0.543)	(0.003)
Player-match fixed effects	Yes	Yes	Yes	Yes
Polynomial throw number (order)	5	4	9	8
Observations	94,040	94,040	94,030	282,110
Adjusted R-squared	0.119	0.117	0.109	0.118
Panel D: Youth				
Both can win match	-1.270	-1.171	-0.680	-1.072^{**}
	(0.112)	(0.141)	(0.399)	(0.020)
Player-match fixed effects	Yes	Yes	Yes	Yes
Polynomial throw number (order)	7	8	9	9
Observations	29,738	29,738	29,724	89,200
Adjusted R-squared	0.129	0.142	0.137	0.146

Notes: The table reports the coefficients resulting from OLS regression analyses of the number of points thrown in the first, second, and/or third throw of a leg. The regression specifications include a polynomial of order n for the player's prior number of throws, where n is chosen to maximize the adjusted R-squared. Other definitions are as in Table 2.

as possible is sub-optimal.[5] In contrast, in subsequent throws a player may try to maneuver herself into a favorable finishing position, for example by trying to reach a remaining score of 40. As in the previous analysis, we control for skill differences between players and through time by including player-match fixed effects, and for possible warming-up and fatigue effects within a match by including a polynomial of order n for the player's number of throws in the match prior to the current throw, where the value of n is chosen to maximize the adjusted R-squared.

Table 3 presents the OLS regression results. The first three models show how players' average number of points is affected if both players can win the match by

[5] Technically, players may already want to aim for less than the maximum number of points in throw three if they did exceptionally well in the first two throws. Such situations are, however, extremely rare. Furthermore, our strategy of using the first three throws and excluding all subsequent throws is empirically supported by the average number of points: this statistic is stable across throw one, two and three, and decreases from throw four onwards.

winning the leg, for the first three throws separately. The coefficients are consistently negative, but not always statistically significant. Model 4 shows the results for the first three throws combined. In each of the four samples, throwing performance significantly deteriorates in the first three throws of decisive legs. The effect is strongest for BICC and weakest for International players. After scaling the coefficients by the average number of points per throw in the first three throws (see Table 1), these players' performance deteriorates by 2.5% and 0.7%, respectively. For Super League and Youth the performance decrease is 1.5% and 1.8%, respectively.

Following our approach in the previous section, we rerun the analyses using International leg-format matches only, to alleviate the concern that set-format matches have diluted the possible evidence of choking among professional players. Table B2 in Appendix B gives the results. The choking effects become slightly more pronounced, but remain substantially weaker than those for the other three samples.

Altogether, the results in this section corroborate the previous finding that a vast increase in incentives can have a perverse effect on the performance of darts players, and that professionals are less susceptible of this choking under pressure than amateur and youth players.

4 Conclusions and Discussion

The present paper examines how naturally-occurring within-match variation in the incentive to perform well impacts performance in darts. The game of darts offers an attractive natural research setting, because performance can be observed at the individual level and without the obscuring effects of risk considerations and the behavior of others. Our results speak to a growing literature on the limits of increasing incentives as a recipe for better performance. Insofar as incentives increase only moderately, our observations support the classical prediction in economics that higher incentives improve performance. Under really high incentives, however, darts players display symptoms of choke under pressure (Baumeister, 1984; Beilock, 2010).

Using four large data sets that cover the matches of different categories of players—ranging from amateur youth players to professional adults—we find that players perform better under moderately higher incentives. At the crucial stage of a leg, when players have to finish by throwing a so-called 'double', they throw more accurately the more the pressure is on: the fewer darts the opponent needs to finish, the more often players successfully hit their double. The improvement is substantial and of a similar magnitude across all samples. For instance, when the opponent needs only one dart, players are 11-16 percentage points more likely to finish as compared to situations where the opponent cannot instantly finish because she needs more than three darts.

Under really high incentives, however, the performance of youth and amateur players diminishes. If both players can win the leg with just one dart, their finishing performance is significantly worse in decisive legs where both can win the entire match than in non-decisive legs: in decisive legs these types of players are 7-13 percentage points less likely to finish than in non-decisive legs. Such choking under pressure also occurs at the start of decisive legs. In the early stage of a leg, the optimal strategy is to throw as many points as possible. When both players can win the match by winning the leg, youth and amateur players throw 1.5-2.5% fewer points in their first three throws than they normally do.

Professional players appear less susceptible of choking under the high pressure of decisive legs than youth and amateur players: there is no compelling evidence of deteriorating finishing performance, and for the number of points thrown the effect size is relatively small. This difference may suggest that choking under pressure can be prevented by training and decreases with experience. At the same time, however, because the ability to deal with pressure is a competitive advantage, the lower sensitivity of professional players may also be the result of a selection effect.

Most of the earlier studies that have found evidence for choking under pressure in sports did exclusively consider professionals. Our analysis uses a broader set of players, and suggests that choking under pressure is of an even bigger concern for lower-skilled individuals. Furthermore, these earlier studies generally suggest that higher pressure always harms performance. The present paper shows that the impact can also be positive, as long as the increase is only moderate.

References

Ariely, D., Gneezy, U., Loewenstein, G. and Mazar, N. (2009), 'Large stakes and big mistakes', *Review of Economic Studies* **76**(2), 451–469.

Baumeister, R. F. (1984), 'Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance', *Journal of Personality and Social Psychology* **46**(3), 610–620.

Baumeister, R. F. and Showers, C. J. (1986), 'A review of paradoxical performance effects: Choking under pressure in sports and mental tests', *European Journal of Social Psychology* **16**(4), 361–383.

Beilock, S. L. (2010), *Choke: What the secrets of the brain reveal about getting it right when you have to*, New York: Free Press.

Camerer, C. F. and Hogarth, R. M. (1999), 'The effects of financial incentives in experiments: A review and capital-labor-production framework', *Journal of Risk and Uncertainty* **19**, 7–42.

Camerer, C. F., Loewenstein, G. and Prelec, D. (2005), 'Neuroeconomics: How neuroscience can inform economics', *Journal of Economic Literature* **43**(1), 9–64.

Cao, Z., Price, J. and Stone, D. F. (2011), 'Performance under pressure in the NBA', *Journal of Sports Economics* **12**(3), 231–252.

Chiappori, P. A., Levitt, S. and Groseclose, T. (2002), 'Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer', *American Economic Review* **92**(4), 1138–1151.

Cohen-Zada, D., Krumer, A., Rosenboim, M. and Shapir, O. M. (2017), 'Choking under pressure and gender: Evidence from professional tennis', *Journal of Economic Psychology* **61**, 176–190.

Dohmen, T. J. (2008), 'Do professionals choke under pressure?', *Journal of Economic Behavior & Organization* **65**(3-4), 636–653.

Gauriot, R. and Page, L. (2015), 'I take care of my own: A field study on how leadership handles conflict between individual and collective incentives', *American Economic Review* **105**(5), 414–419.

Gneezy, U. and Rustichini, A. (2000), 'Pay enough or don't pay at all', *Quarterly Journal of Economics* **115**(3), 791–810.

Hahn, J. and Newey, W. (2004), 'Jackknife and analytical bias reduction for nonlinear panel models', *Econometrica* **72**(4), 1295–1319.

Harb-Wu, K. and Krumer, A. (2017), Choking under pressure in front of a supportive audience: Evidence from professional biathlon. Working Paper.

Hickman, D. C. and Metz, N. E. (2015), 'The impact of pressure on performance: Evidence from the PGA tour', *Journal of Economic Behavior & Organization* **116**, 319–330.

Langer, E. J. and Imber, L. G. (1979), 'When practice makes imperfect: The debilitating effects of overlearning', *Journal of Personality and Social Psychology* **37**(11), 2014–2024.

Lazear, E. P. (2000), 'Performance pay and productivity', *American Economic Review* **90**(5), 1346–1361.

Levitt, S. D., List, J. A., Neckermann, S. and Sadoff, S. (2016), 'The behavioralist goes to school: Leveraging behavioral economics to improve educational performance', *American Economic Journal: Economic Policy* **8**(4), 183–219.

Lindner, F. (2017), Choking under pressure of top performers: Evidence from biathlon competitions. Working Paper.

Paarsch, H. and Shearer, B. S. (1999), 'The response of worker effort to piece rates: Evidence from the British Columbia tree-planting industry', *Journal of Human Resources* **34**(4), 643–667.

Paserman, M. D. (2010), Gender differences in performance in competitive environments: Evidence from professional tennis players. Working Paper.

Price, J. and Wolfers, J. (2010), 'Racial discrimination among NBA referees', *Quarterly Journal of Economics* **125**(4), 1859–1887.

Shearer, B. (2004), 'Piece rates, fixed wages and incentives: Evidence from a field experiment', *Review of Economic Studies* **71**(2), 513–534.

Smith, V. L. and Walker, J. M. (1993), 'Monetary rewards and decision cost in experimental economics', *Economic Inquiry* **31**(2), 245–261.

Toma, M. (2017), 'Missed shots at the free-throw line', *Journal of Sports Economics* **18**(6), 185–207.

Yerkes, R. M. and Dodson, J. D. (1908), 'The relationship of strength of stimulus to rapidity of habit formation', *Journal of Comparative Neurology and Psychology* **18**, 459–482.

Appendix

A Points-Left Fixed Effects

Table A1: Regression results for the likelihood of finishing, controlling for points-left fixed effects

	Model 1	Model 2	Model 3
Panel A: International			
Opp. can finish with 1 dart	0.074***		0.076***
	(0.000)		(0.000)
Opp. can finish with 2 darts	0.048***		0.049***
	(0.000)		(0.000)
Opp. can finish with 3 darts	0.016*		0.016*
	(0.074)		(0.083)
Both can win match		−0.027**	−0.002
		(0.035)	(0.970)
Opp. can finish with 1 dart × Both can win match			−0.037
			(0.392)
Opp. can finish with 2 darts × Both can win match			−0.019
			(0.661)
Opp. can finish with 3 darts × Both can win match			0.003
			(0.951)
Player-match fixed effects	Yes	Yes	Yes
Points-left fixed effects	Yes	Yes	Yes
Polynomial throw number (order)	3	3	3
McFadden pseudo R-squared	0.363	0.361	0.363
Observations	111,241	111,241	111,241
Panel B: BICC			
Opp. can finish with 1 dart	0.082***		0.091***
	(0.000)		(0.000)
Opp. can finish with 2 darts	0.051***		0.056***
	(0.000)		(0.000)
Opp. can finish with 3 darts	0.027***		0.031***
	(0.000)		(0.000)
Both can win match		−0.093***	−0.003
		(0.000)	(0.805)
Opp. can finish with 1 dart × Both can win match			−0.123***
			(0.000)
Opp. can finish with 2 darts × Both can win match			−0.071***
			(0.000)
Opp. can finish with 3 darts × Both can win match			−0.052***
			(0.001)
Player-match fixed effects	Yes	Yes	Yes
Points-left fixed effects	Yes	Yes	Yes
Polynomial throw number (order)	5	5	5
McFadden pseudo R-squared	0.331	0.330	0.332
Observations	587,876	587,876	587,876

Notes: The table reports the average marginal effects resulting from logit regression analyses of finishing performance across all throws where the player can finish the leg with one dart. The regression specifications now in addition include points-left fixed effects. Definitions are as in Table 2.

Table A1: Continued

	Model 1	Model 2	Model 3
Panel C: Super League			
Opp. can finish with 1 dart	0.080^{***}		0.086^{***}
	(0.000)		(0.000)
Opp. can finish with 2 darts	0.049^{***}		0.051^{***}
	(0.000)		(0.000)
Opp. can finish with 3 darts	0.025^{**}		0.028^{**}
	(0.020)		(0.012)
Both can win match		-0.040^{***}	0.032
		(0.003)	(0.403)
Opp. can finish with 1 dart × Both can win match			-0.102^{**}
			(0.013)
Opp. can finish with 2 darts × Both can win match			−0.049
			(0.235)
Opp. can finish with 3 darts × Both can win match			−0.056
			(0.225)
Player-match fixed effects	Yes	Yes	Yes
Points-left fixed effects	Yes	Yes	Yes
Polynomial throw number (order)	5	5	5
McFadden pseudo R-squared	0.324	0.323	0.324
Observations	73,191	73,191	73,191
Panel D: Youth			
Opp. can finish with 1 dart	0.086^{***}		0.102^{***}
	(0.000)		(0.000)
Opp. can finish with 2 darts	0.028^{*}		0.032^{**}
	(0.064)		(0.042)
Opp. can finish with 3 darts	0.012		0.012
	(0.436)		(0.485)
Both can win match		-0.073^{***}	−0.007
		(0.000)	(0.866)
Opp. can finish with 1 dart × Both can win match			-0.115^{**}
			(0.011)
Opp. can finish with 2 darts × Both can win match			−0.036
			(0.441)
Opp. can finish with 3 darts × Both can win match			0.010
			(0.842)
Player-match fixed effects	Yes	Yes	Yes
Points-left fixed effects	Yes	Yes	Yes
Polynomial throw number (order)	5	5	5
McFadden pseudo R-squared	0.343	0.342	0.345
Observations	33,454	33,454	33,454

B Leg-Format International Matches Only

Table B1: Regression results for likelihood of finishing, leg-format International matches only

	Dependent variable:		
	Finishing		
	(1)	(2)	(3)
Opp. can finish with 1 dart	0.108^{***}		0.110^{***}
	(0.000)		(0.000)
Opp. can finish with 2 darts	0.072^{***}		0.073^{***}
	(0.000)		(0.000)
Opp. can finish with 3 darts	0.031^{***}		0.031^{***}
	(0.001)		(0.001)
Both can win match		-0.026^{*}	0.010
		(0.057)	(0.817)
Opp. can finish with 1 dart × Both can win match			-0.049
			(0.289)
Opp. can finish with 2 darts × Both can win match			-0.028
			(0.553)
Opp. can finish with 3 darts × Both can win match			0.002
			(0.963)
Player-match fixed effects	Yes	Yes	Yes
Polynomial throw number (order)	3	3	3
McFadden pseudo R-squared	0.331	0.328	0.331
Observations	100,292	100,292	100,292

Notes: The table reports the average marginal effects resulting from logit regression analyses of finishing performance across all throws where the player can finish the leg with one dart, for International matches with a leg format only. Definitions are as in Table 2.

Table B2: Regression results for the number of points thrown, leg-format International matches only

	Model 1: Throw 1	Model 2: Throw 2	Model 3: Throw 3	Model 4: Throws 1-3
Both can win match	-0.277	-0.803	-1.077^{*}	-0.721^{**}
	(0.650)	(0.191)	(0.077)	(0.040)
Player-match fixed effects	Yes	Yes	Yes	Yes
Polynomial throw number (order)	7	4	3	9
Observations	142,838	142,838	142,831	428,507
Adjusted R-squared	0.240	0.242	0.236	0.250

Notes: The table reports the coefficients resulting from OLS regression analyses of the number of points thrown in the first, second, and/or third throw of a leg, for International matches with a leg format only. Definitions are as in Table 3.