Performance When It Matters
Psychological Pressure in a Natural Setting

By Rik Boijmans

Erasmus School of Economics

Abstract

Pressure is expected to increase effort and motivation, factors that increase performance, while at the same time pressure can diminish performance when there is too much of it. For a performance to be maximized it should be understood how people react to different moments of pressure. By analyzing the performance of volleyball players under different pressure related situations this paper provides empirical evidence from a natural setting coherent with a high external validity. Loss aversion theory is considered to be a determent factor for changing pressure situations. Performance is then influenced by the change in pressure. Not just in sports, but in many aspects of life pressure plays a role and understanding the effects on performance can help to manage these performances.

Keywords: Performance under pressure, choking under pressure, loss aversion theory.

Introduction

Moments of pressure can be observed in many different situations of life like exams, presentations, deadlines or any other moments in life when we feel either self-imposed or peer-imposed pressure. Despite its ubiquity in everyday life, there are surprisingly few studies that empirically examine the effects of psychological pressure on performance outside behavioral labs. While laboratory experiments guarantee internal validity and help us establish the causal effects of psychological pressure on variables such as performance, some question whether lab manipulations of psychological pressure are sufficiently close to the real-world pressure mechanisms. Thus, it is important to gather field evidence to establish also the external validity of these findings. This is the goal of my thesis.

Psychological research during the 1970s did not review pressure as such. Instead, at that time, psychologists tended to consider pressure as any mechanism that would increase one's attention and dedication to a particular task, and would therefore increase performance. Baumeister (1984) was one of the first to identify that increased attention does not always lead to a better performance. Ariely, Gneezy, Loewenstein and Mazar (2009) reviewed the two principles regarding extrinsic motivation vs. performance: (i) increasing performance-related incentives will result in a higher motivation and effort, and (ii) increasing motivation and effort will lead to a better performance. Only taken into consideration extrinsic motivators, the first assumption regarding higher incentives is generally accepted.[1] Experiments regarding the second assumption, though originated by incentives, show that performance follows an inverted U pattern. This pattern was

[1] Ariely, Gneezy, Loewenstein and Mazar (2009) state this themselves, but do mention some literature that objects to this theory like Camerer, Babcock, Loewenstein, & Thaler (1997) regarding cab drivers quitting a day's work early when hourly earnings were high. In addition, Amabile (1993) argues that extrinsic motivation often decreases intrinsic motivation, where the latter is regarded as more powerful than the former. This topic – intrinsic motivation – is though not further discussed in this paper.

first introduced by Yerkes & Dodson (1908) for which Baumeister (1984) introduced the term ‘choking’ which refers to a decline in performance where extrinsic motivators increase.

Still, different views exist because most work done on the subject used an experimental setting to examine the effects. Opposing views argue that the external validity is low and therefore question the applicability in real life (see Apesteguia & Palcios-Huerta (2010) and Cao, Price, & Stone (2011)). In order to conclude this debate, the effects of psychological pressure on performance in the field need to be examined. The first step is to find a real life situation in which only the pressure component varies (e.g. over time) in a situation that is the same in all other aspects. In other words, I needed to find a situation where psychological pressure varies naturally in an otherwise “ceteris paribus” condition. The difficulty is that in nature, perfectly comparable ceteris paribus conditions are not easy to find as, typically, there are multiple competing factors that need to be accounted for. This makes a real life situation complex to analyze due to the existence of unobservable correlates that may confound the effect one is trying to measure.

In this paper, I focus on the effect of psychological pressure in sports in an attempt to account for external validity, while at the same time examining behavior in a highly controlled environment. The advantage of sports is that situations are restricted by rules and the goal is always clear; to win. Still, almost all sports, or situations within sports, are dynamic and contain multiple factors to be accounted for.

Other studies also addressed external validity and used sports data account for this issue. Apesteguia and Palcios-Huerta (2010) assessed penalty shoot-outs in soccer and Cao, Price and Stone (2011) evaluated free throws in the NBA (basketball). Both argue that pressure is a significant factor that may influence athletes’ and teams’ performances. With the penalty shoot-outs Apesteguia and Palcios-Huerta argue that the second kicker is more under pressure as a result of an ahead-behind asymmetry. For basketball, the free throws closer to the end of a match are regarded as pressure ‘free throws’. The sports related studies mentioned in this paper use static sporting moments. A *static* sporting moment is considered to be an action that does not require any quick anticipation of the game situation at hand. Some good examples are free throws in basketball (always the same distance, height, ball etc.), darts but also a serve in volleyball or tennis. All actions start in the exact same setting, without any in-game actions to alter this.[2] I will refer to non-static moments as *dynamic* moments. The above sports related studies argue that due the focus on such static moments the pressure aspect of the act is one of the few remaining differentiating variables.

I argue that the problem with these studies is that static moments are only present in some sports (e.g. golf, darts etc.) and occur very infrequently in real life. Therefore, they do not relate to the in-game moments present in many other sports – and many other life situations – which may contain a very different and more complex pressure experience. Furthermore, prior studies, have typically ignored the role of the opponent (e.g. free throws in basketball). The absence of opposition[3] within static moments makes these situations unrealistic and not very representative of more general high-pressure tasks. The combination of no opposition within static moments eliminates parts of the cognitive performance, i.e. quickly making the right decision or anticipating on a situation. An important contribution of this paper is to assess ‘in-game’ pressure in a more realistic and complex environment. I argue that that the presence of an opposing team enforces the effect of cognitive pressure.

[2] An opponent can have some effect on a serve in e.g. volleyball or tennis by positioning differently, but the serve itself, whether placed somewhere else as a result of the opposition, should not change in itself. To elaborate, all serves to the left, whether forced by the opposition, take place in the same setting and should be of the same quality.

[3] For soccer penalties there is a goalkeeper, but the rule of thumb is that a well taken penalty should always score. Opposition is therefore limited.

I also argue, that it is fairly well shown that pressure influences *cognitive* performance, but that only few studies have looked into the effect of psychological pressure on *physical* performance in combination with a cognitive aspect. For instance, in tennis, the *cognitive* aspect would be: “which way do I hit the ball?”, followed by the *physical* performance, or ability, of hitting the ball as planned.[4] Both aspects influence the quality of the overall performance, for which I argue, that this resembles a real life situation the most and has been infrequently covered in the on-topic papers.

In line with the previous statements an analysis on volleyball data suits the topic of this paper. I will use data from the Dutch male and female national volleyball competition to examine performance as a result of varying pressure moments. Volleyball itself is a perfect setting in the sense that every game action is closely related to a point (i.e. performance) though still *dynamic*.[5] Specifically, volleyball is a rally sport, that is, after every point players can contemplate on their current situation and consciously or unconsciously realize the pressure they feel. The difference from e.g. penalty kicks, is that after a serve which starts the rally, players have little time to actively regard or recall the pressure they are under, which is precisely how the *dynamic* sport condition differs from *static* sporting moments in relation to pressure.

The data I will use is data of the Dutch national league in the season 2014/2015 for both men and women. During the regular season a total of 136.508 ball touches were recorded. I will determine the so called ‘big points’ (which, in volleyball, is a common term for pressure points) and I will show how the enhanced pressure influences scoring probabilities.

[4] Note that the major difference between static and dynamic moments is the cognitive time to make a decision and the time between a final decision and the performance.

[5] With the exception of the serve.

I propose a theory-based account of why psychological pressure influences performance. Following Apesteguia & Palcios-Huerta (2010), I will take into account the presence of the ‘ahead-behind’ asymmetry in volleyball, which I will state, has a resemblance with the general accepted theory of loss aversion.

This paper is structured as follows. The first section will review the relevant literature (I. Theory). In the second section I will describe how the volleyball game serves as a natural experiment and how the data is structured (II. Data description). In the third section I describe my methodology (III. Research methodology). In the fourth section I will present the results of my analyses (IV. Analysis & Results). For the second and forth section a better understanding of the volleyball game can be helpful, therefore – for those unfamiliar with volleyball - I added a simple appendix describing the volleyball game. In the fifth section I will discuss the implications of my results and opt some areas for further research (V. Conclusions).

I. Theory

In this paper, I build on two key theoretical streams of literature: (i) theories on pressure (vs. performance) (see, Baumeister (1984) and Beilock, Kulp, Holt, and Carr (2004)) and (ii) loss aversion theory (see, Kahneman & Tversky (1979)). In the introduction of this paper pressure vs. performance is briefly discussed. The loss aversion phenomenon was briefly mentioned in relation to the 'ahead-behind' asymmetry. This section will discuss both topics extensively.

Psychological Pressure Theories

Psychologists have looked primarily to the effects of psychological pressure on cognitive performance. For instance, scholars have studied subjects solving math problems to examine performance under different levels of psychological pressure (see, Beilock & Carr (2005) and Markman, Todd Maddox & Worthy (2006)). These studies tend to show that cognitive performance is suppressed when under pressure. Beilock & Carr (2005) refer to three issues regarding the influence of pressure on math performance, i.e. anxiety, stereotype threat and choking under pressure. The last one, choking under pressure, is found to be a mental distraction resulting in a performance less up to par when additional pressure is present (Beilock, Kulp, Holt, & Carr, 2004). Anxiety and a stereotype threat are not subjects in this paper. The focus will therefore be on choking under pressure.

Where the majority of research regarding choking under pressure (pressure vs. performance), covers cognitive performance, some studies tried to fit physical performance with pressure through sport settings. Apesteguia & Palcios-Huerta (2010) and Cao, Price, & Stone (2011) found a significant difference in sports performance with regard to their determined pressure moments. Both papers use expert performances, including top sporting leagues for empirical research.

Apesteguia & Palcios-Huerta (2010) used penalty kicks in football (soccer) during major tournaments for their research. They found a significicint difference in winning probabilities between teams that go first and teams that follow second. They argue that the lagging team in the shoot-out feels the additional pressure of an ‘ahead-behind’ asymmetry, of which I will refer to as the effect of loss aversion (see next paragraph). Kocher, Lenz & Sutter (2012) also assessed penalty shoot-outs but found no significant difference from a 50/50 chance of winning.

Cao, Price, & Stone (2011) used free throws in basketball for their analysis of performance under pressure. They assessed free throw data of eight NBA seasons to see if there is a significant difference between the probability of scoring a free throw under (additional) pressure or without (additional) pressure. They found a significant difference in the success-rate of free throws with additional pressure, that is, free throws at the end of a game where the score is close to a tie.

Francesco, Innocenti and Pin (2013) argued that all chosen sporting moments mentioned above, though externally valid, include large stakes. Therefore Francesco, Innocenti and Pin conducted an experiment where the stakes were not that high and found that individuals did behave differently in different pressure situations. It is argued though that pressure was too low in all situations, which makes performance under pressure not the best matching dependent variable to base conclusions on.

Loss Aversion Theory

Kahneman & Tversky (1979) introduced loss aversion theory which is in line with the 'ahead-behind' asymmetry. Loss aversion theory states that losses resonate more than gains of the same magnitude for which I argue that the term Apesteguia & Palcios-Huerta (2010) used, 'ahead-behind' asymmetry, comes from loss aversion and is the variable that influences winning probabilities under otherwise 'ceteris paribus' situations.

Loss aversion theory explains the psychological effect of a loss (disutility) of the same magnitude is valued as larger than a gain (utility) of the same magnitude. This theory is something that marketers have become very handy with in influencing people's behavior. Trial periods are an example of letting consumers experience a service, where at the point that the period ends, this loss of the service is experienced as a larger disutility than the utility that is expected to gain before the trial period. Thus, stimulating sales.

In sports, loss aversion theory comes into play when a point, rally or act is closely related to the moment of winning or losing (a part of) a match. First of all, the difference between the winning utility and the losing disutility comes in sight (i.e.: winning utility > losing disutility), resulting in that pressure tends to grow to which the phenomenon of chocking under pressure is related (Baumeister, 1984). Secondly, when the difference comes in sight between the winning utility and the losing disutility, loss aversion theory suggests that:

$$|\textit{winning utility}| < |\textit{losing disutility}|$$

In other words, the stakes of losing – the absolute magnitude of the losing disutility – are higher than the absolute winning utility and I therefore argue that this results into more pressure. Furthermore I argue that the team that relates a point most with the losing disutility is more likely to 'choke under pressure', and thus performs worse. In other words, does a team's state of mind focusses on a possible loss or a possible win? To confirm a difference in pressure and review a team’s state of mind I will review the pressure points within the volleyball game.

Within the volleyball game a term often used is ‘big points’. Volleyball athletes use the term ′big points’ to refer to the in-game points that are known or meant to trigger a higher level of psychological pressure among players, as compared to other in-game points. In volleyball these ‘big points’ occur typically close to the end of a set (i.e.: points are closely related to the success of a set) and when the score is close to a tie (i.e.: within a few points difference). I will mention the exact specifications in the data section. In this paper I will refer to these ‘big points’ as ‘pressure points’.

To be able to test for loss aversion, I classify pressure points as either negative pressure points, positive pressure points, or equal pressure points. Negative pressure points are pressure points that occur when the referred (or focal) team is lagging behind in the in-game score. Positive pressure points occur when the referred (or focal) team is ahead in the in-game score. Equal pressure points are those where the score is tied.

After determining the pressure points, I am able to test whether (i) pressure points do influence a team’s performance and if so (ii) how the different pressure points (negative, positive or equal) influence this performance.

For a manager – whether this is in sports or not – it is something to be conscious and cautious about when making decisions in such circumstances. I will discuss on the findings and implications as such in chapter five.

II. Data description

To test whether psychological pressure has an influence on physical performance, I analyzed a large dataset with the results of 308 volleyball games. This data is recorded by several Dutch volleyball scouts who – together with the Dutch national volleyball team – have the goal to track and analyze the game. For this paper I used the complete data set of the 2014/2015 Nevobo A-league regular volleyball competition, both male and female leagues. Every ball contact is logged in this dataset, which I will refer to as ‘actions’ throughout this paper. Every action is logged with an according type (serve, attack etc.) and efficacy (how well the action is performed). Table 1 shows number of actions present in the dataset. The rows indicate the gender and the type of actions. The columns in table 1 refer to the efficacy scores, which are further described in table 2. The numbers in table 1 indicate the number of actions with according efficacy scores that are present in the dataset. This means that within the dataset e.g. 3,150 male serves are logged with an efficacy score of ‘=’. In total, 20,591 male serves are logged. The total amount of actions initially used for this paper is 123,901.

Although several scouts have logged the data, guidelines have been extensively discussed so that all efficacy ratings are aligned as best as possible. I have only used the data of the regular season which means every team played each other twice, so that opposition is as similar as possible for every team.

Table 1 - Description of the Dataset

Gender	Action	=	/	-	!	+	#	Total
Male	S(erve)	3,150	742	6,821	4,074	4,419	1,385	20,591
	R(eception)	1,283	740	3,131	1,615	4,745	5,227	16,741
	A(ttack)	2,362	2,068	1,803	306	6,504	10,697	23,740
	B(lock)	1,840	8	139	195	596	2,060	4,838
	F(ree ball)	21	4	45	9	472	186	737
	E (set)	447	8	5	1	1,485	2	1,948
	D(ig)	784	181	178	131	960	87	2,321
Total		9,887	3,751	12,122	6,331	19,181	19,644	70,916
Female	S(erve)	1,708	675	3,558	2,949	3,407	1,517	13,814
	R(eception)	1,493	646	2,821	1,076	2,538	3,581	12,155
	A(ttack)	1,719	1,411	1,905	850	5,308	6,743	17,936
	B(lock)	692	538	126	379	256	1,451	3,442
	F(ree ball)	50	20	51	9	443	635	1,208
	E (set)	439	2	5	-	1,525	-	1,971
	D(ig)	685	251	484	250	553	236	2,459
Total		6,786	3,543	8,950	5,513	14,030	14,163	52,985
Total		**16.673**	**7.294**	**21.072**	**11.844**	**33.211**	**33.807**	**123.901**

The efficacy variable is valued on a six point scale (see table 2 for an indication on efficacy scales). Although the same scale is used across all actions, every action (serve, attack etc.) will have a different meaning per efficacy scale. As table 2 states, a ‘#’ means a perfect action. For an attack, a perfect action means that a point is scored directly, while for a reception a perfect action means that the ball is played perfectly into the hands of the setter. Furthermore, the efficacy of the serves and receptions is related to one another. A good serve means a bad reception, and vice versa. Yet, for the purposes of statistical inference in my thesis, the important thing to keep in mind is that – despite their qualitative differences – in terms of ordinal properties, ‘/’ is always better than ‘=’ and ‘#’ is better than ‘+’ and so on.

Table 2 - Efficacy score scale (actions)

=	Direct point for opposition by fault
/	Bad action (very easy serve/attack, bad reception)
-	Weak action (easy serve/attack, weak reception)
!	Decent action (decent serve/attack, decent reception)
+	Good action (good serve/attack, good reception)
#	Perfect action (scoring serve/attack, perfect reception)

In a volleyball game, two teams play against each other a series of points. For each point, the teams play a series of actions, (such as serve, reception, attack, etc.) to determine who wins the point. A point is won by a fault of the opponent (a '=' efficacy) or by a scoring action (a '#' efficacy for either an attack, serve or block). I was able to construct a dependent variable by first identifying all points, with according actions in the data set. Using the efficacy scale in the database, I developed an ordinal scale to indicate the quality of these points. This scale refers to how well a point is played. All 123,901 actions cumulate to a total of 34.416 points of which 32.806 were useable and valued on the ordinal scale. For the 1.610 excluded cases it was not possible to appoint a quality scale measurement due to an unclear sequence and/or efficacy score of actions.

Table 3 shows the meaning per classification on the ordinal scale. Do note that the scale is from the perspective of the serving team. For a serving team, a point quality of '2' is better than '1', a point quality of '4' is better than '3', and so forth. For a receiving team though, an ace or direct score on serve is the *worst* case scenario because a point is directly lost. As table 3 indicates, a direct score on serve should though be valued with a '6'. A '6' indicates that this is the *best* scenario for a point which is not the case from a receiving team perspective. In the receiving ordinal scale column the ordinal scale is reversed for the perspective of a receiving team, so that in all cases (whether the team served or received) a '1' is the worst outcome and a '6' is the best possible outcome. The first two columns denote the difference between a team's perspective. Furthermore, the two columns on the right in table 3 indicate whether a point is won or lost. Below table 3 the descriptions are elaborated further.

Table 3 - Ordinal point quality description[6]

Ordinal Scale[7]		No.	Point description	Point won or lost	
On serve	Receiving			On serve	Receiving
1	6	4,885	Fault serve	Point lost	Point won
2	5	10,879	Receiving team scores a point with fewer than 5 actions (but more than 1)	Point lost	Point won
3	4	3,934	Receiving team scores a point with 5 actions or more	Point lost	Point won
4	3	2,135	Serving team scores a point with 6 actions or more	Point won	Point lost
5	2	8,084	Serving team scores a point with fewer than 6 actions (but not scale condition 6)	Point won	Point lost
6	1	2,889	Ace or direct score on serve	Point won	Point lost

1. Fault serve

If an action is a serve and the efficacy score is '=' this constitutes as a fault serve and is valued with a 1. (No receiving team action is logged which means no receiving team's action is valued with a quality point score of a 6.)

2. Receiving team scores a point with fewer than 5 actions (but more than 1)

The point is scored by the receiving team and the number of actions in the point are less than 5, then the point is valued with a 2. (For a receiving team a point with these conditions is valued with a 5.)

3. Receiving team scores a point with 5 actions or more

All points scored by the receiving team with more than 4 actions are valued with a 3. (For a receiving team a point with these conditions is valued with a 4.)

[6] In the appendix I will elaborate on the fact why a '5' is considered to be better than a '4' which includes some elaboration on the volleyball game. I will furthermore elaborate on how I constructed this data.

[7] It should be noted that in the dataset every single action is valued with an ordinal scale on the quality of the point. This results in every action being valued as an action that contributes to a specific quality of that point.

4. Serving team scores a point with 6 actions or more

All points scored by the serving team in more than 5 actions are valued with a 4. (For a receiving team a point with these conditions is valued with a 3.)

5. Serving team scores a point with fewer than 6 actions (but not scale condition 6)

All points scored by the serving team in less than 6 actions, excluding aces or direct scores from a serve. (For a receiving team a point with these conditions is valued with a 2.)

6. Ace or direct score on serve

All points scored by the serving team due to a result of a quality serve. That is, an ace (serve has a '#' as efficacy score) or a serve with a failing reception (reception after the serve has a '=' as efficacy score). (For a receiving team a point with these conditions is valued with a 1.)

Table 4 shows the distribution of points regarding this ordinal point quality scale. The amount of points in the dataset is shown per gender and if a point is being played a under (additional) pressure. Below table 4 I elaborate on when a point is played under pressure or not and is considered a pressure point.

Table 4 - Ordinal point indication

	Gender	1	2	3	4	5	6	**Total**
No pressure points	Male	2,867	6,407	1,843	917	4,254	1,283	17,571
	Female	1,601	3,387	1,641	1,051	3,087	1,394	12,161
Pressure points	Male	295	780	297	97	503	118	2,090
	Female	122	305	153	70	240	94	984
Total		4,885	10,879	3,934	2,135	8,084	2,889	32,806

Pressure points

I constructed an ‘in-game score’ from the original data in the dataset which did not include this variable.[8] This in-game score refers to the in-game score at the time a specific point is played, e.g. whether a point is played at an in-game score of ‘2-2’ or at ‘23-23’. The in-game score is then used to determine which points are regarded as pressure points. A pressure point[9] is considered to be a point when teams are more likely to feel additional pressure as a result of the in-game score. In this paper a point is considered to be a pressure point when one or both teams are above 18 points (8 points when it is a fifth set[10]) and the difference between the team’s scores is four or less. If a point is a pressure point it is classified as either a negative-, positive-, or equal pressure point, as discussed in the theoretical section. All points that do not comply with these conditions are regarded as ‘No pressure point’.

All points are identified and valued on the ordinal point quality scale and as either a pressure point or not. As discussed before, a point consists of multiple actions, whereas an action is a single act of 1 player. Analysis will take place at the action level. This means that the pressure point and ordinal point quality scale should be included as a variable for every action. Table 5 shows an example of a piece of the updated dataset with the added variables, denoted with an ‘*’. In this example a serve with a ‘+’ efficacy under a positive pressure point condition results in a ordinal point quality scale of 5[11], being the dependent variable.

[8] Notice that the ‘efficacy score’ refers to the scores in table 2 and that the ‘in-game score’ refers to the current set score a point is played.

[9] I.e. big points in volleyball terminology.

[10] A volleyball set is won when a team reaches 25 points with a two point difference. Only a fifth set is played up to 15 points. Every individual copes differently with pressure and will have different definitions about a pressure point. Some might feel added pressure at a score of ‘15-15’ while others do not feel any pressure at a score of ‘20-20’. An acceptable average is sought and found in the restrictions described in the main text.

[11] The point with point ID 5 is won by the serving team in 3 actions (less than 6). The opposition makes a fault in the attack, denoted by the ‘=’ efficacy.

Table 5 - Updated dataset

Action ID	Point ID*	Team ID	Type of action	Efficacy	Pressure point*	Ordinal point quality scale*
15	5	1	Serve	+	Positive	5
16	5	2	Reception	#	Negative	2
17	5	2	Attack	=	Negative	2

Table 6 shows the distribution of pressure point actions that are used in this research. It e.g. shows that 63,176 actions are present in the dataset that are played by a male under a ‘no pressure point’ condition. A total of 4,688 negative pressure point actions, 4,599 positive pressure point actions and 2,425 equal pressure point actions are logged in the dataset.

Table 6 - Pressure point actions

	Male	Female	Total
No pressure point	63,176	49,013	112,189
Neg. pressure point	3,048	1,640	4,688
Pos. pressure point	2,950	1,649	4,599
Eq. pressure point	1,742	683	2,425

III. Research methodology

A considerable disadvantage of using ordinal data as the independent variable is that the distance between values is not expected to be the same over the entire scale. A model that accounts for this problem is the Ordered Probit model.

$$y^* = x'\beta + u \qquad u \div \sim N(0,1)$$

The Ordered Probit model measures the underlying latent ‘y*’ with an utility function in which ‘u’ stands for a normal distribution with a mean of 0 and a standard deviation of 1, thus a standard normal distribution.

For M being the amount of ordinal values, M-1 thresholds are identified. When these thresholds are determined, the latent variable is used as mean for a normal distribution. The surface under the normal distribution graph between the thresholds is equivalent to the probability of the ordinal value allocated to these intervals. With M-1 thresholds leaves M surfaces under the normal distribution graph.

In figures 1 and 2 all surfaces are denoted by ‘y=1’, ‘y=2’, up to ‘y=6’. In figure 1, a situation is drawn where $x'\beta$ is 0. Figure 2 resembles a situation where $x'\beta$ is equal to -1. The modelled latent variable – y* – changes the position of the graph, so that the surface of e.g. y=1 is larger in figure 2 than in figure 1. The surface under the graph of a standard distribution is 1, so that the sum of partial surfaces equals to 1. As already mentioned, the surface of e.g. y=1, equals to the probability of the related ordinal scale, described in the data section.

Figure 1 - Normal distribution (x'β = 0)

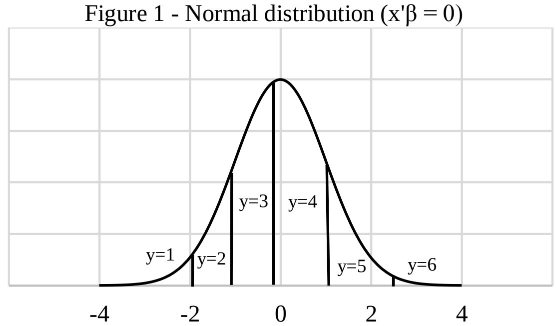
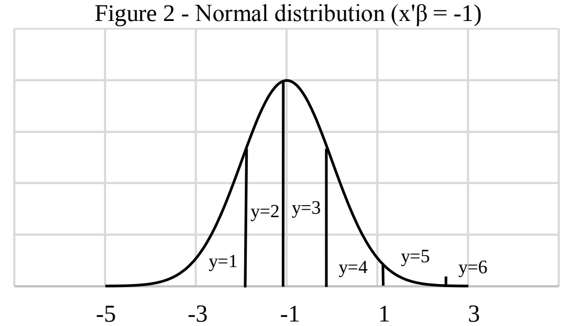


Figure 2 - Normal distribution (x'β = -1)



The functions below denote how to probabilities per y are calculated. In these functions the link between the latent variable – y* – and the observed variable – y – is made.

$$P[y=1] = P[y^* \leq k_1] = \Phi(k_1 - x'\beta)$$
$$P[y=2] = P[k_1 < y^* \leq k_2] = P[y^* \leq k_2] - P[y^* \leq k_1] = \Phi(k_2 - x'\beta) - \Phi(k_1 - x'\beta)$$
$$P[y=3] = P[k_2 < y^* \leq k_3] = P[y^* \leq k_3] - P[y^* \leq k_2] = \Phi(k_3 - x'\beta) - \Phi(k_2 - x'\beta)$$
$$P[y=4] = P[k_3 < y^* \leq k_4] = P[y^* \leq k_4] - P[y^* \leq k_3] = \Phi(k_4 - x'\beta) - \Phi(k_3 - x'\beta)$$
$$P[y=5] = P[k_4 < y^* \leq k_5] = P[y^* \leq k_5] - P[y^* \leq k_4] = \Phi(k_5 - x'\beta) - \Phi(k_4 - x'\beta)$$
$$P[y=6] = 1 - P[y=1] - P[y=2] - P[y=3] - P[y=4] - P[y=5] = 1 - P[y^* \leq k_5] = 1 - \Phi(k_5 - x'\beta)$$

The goal is to model the latent variable with volleyball specific variables so that performance can be predicted through the Ordered Probit model. In addition – to capture the goal of this research – the pressure point variables will be added to see if and how performance is influenced. The negative, positive and equal pressure point dummy variables will reveal how teams react to these situations.

After constructing an Ordered Probit model for the overall volleyball game, another Ordered Probit model is constructed, including the different teams as an interaction term for the different pressure points. This enables us to get insights into how different teams perform under different pressure situations.

IV. Analysis & Results

In this chapter I will start by explaining the models and their chosen configuration. I will support this by explaining related volleyball differences. I will finish this chapter by discussing the implications of the results, where I will discuss the previously stated research question in relation to the findings.

The analysis of the Ordered Probit models is shown in table 7. For the three major actions[12] – that are serve, reception and attack – a separate Ordered Probit model has been run with a distinction between male and female.[13] This results in table 7 showing six different Ordered Probit models, one per column. For all the models, the ordinal point quality scale, that qualifies a point as either good or bad, is the dependent variable.[14] The models consist of two independent variables, the ordinal logged efficacy score and the nominal pressure point variable. The models therefore predict what the probability is of e.g. winning a point in 6 actions or more – P [y = 3] – after e.g. a serve efficacy of '!' during a negative pressure point.

All models review single actions related to the column title. For the male serve model, 20,591 male serves are analyzed, where for the male reception model, 16,740 male receptions are analyzed, and so forth. The efficacy scores should be controlled for when predicting whether a point will be scored or not. A failed attack after a horrible reception under pressure should be denoted as such. Therefore, the efficacy scores are not considered as dependent variables but as independent.

[12] These major actions have been consistently logged by all different scouts and are therefore the best variables for this analysis.

[13] The male and female volleyball game is different in many aspects of the game. To control for this variable in one model would mean that all variables should interact with this dummy variable. Due to the large data set, a separate analysis is possible and preferred.

[14] In general, the ordinal point indication is referred to as performance, though it should be noted that the ordinal point indication regards the relative performance compared to the opposition.

The nominal pressure point variable indicates under what pressure an action took place, no pressure being the base variable. The efficacy score variable indicates the quality of a particular action, so that in the serve models, only serve efficacies are included. This eliminates the problem of similar efficacies having different meaning for different actions. As it is shown in table 7, all efficacy scores are found to be significant.

Table 7 also displays threshold values introduced in the methodology section. For all models M-1 thresholds is correct. I will elaborate per specific model why M is not the same for every model.

The row in table 7 with ‘N actions’ denotes the number of actions are included in the model. The adjusted R^2 values are also displayed in table 7. The higher the adjusted R^2 value, the more the variance of the dependent variable is explained by the independent variables. I will start by elaborating on the different models, after which I will discuss the findings and their implications.

(i) Serve models

The serve models have the highest adjusted R^2 values, 0.718 for the male model and 0.755 for the female model. A big factor in these high adjusted R^2 values comes from the fact that a serve logged with an ‘=’ directly indicates a fault serve. A serve logged with an efficacy score of ‘#’ directly indicates an ace. Both serve efficacies are one-on-one in line with two of the six independent variable outcomes. These parameters are therefore highly contributing to the extent in which the variance is explained by these set of variables. They are included in the model with the main purpose of covering the entire volleyball game. It should be taken into consideration that the adjusted R^2 value is no direct indication that pressure has a high explanatory value when it comes to the variance.

The serve models differ from the other models with regards to the '/' efficacy score parameters. Where the rows are ordered accordingly to its ordinal scale, the '/' stands out. After looking at the data again, it shows that serves are logged with a '/' whenever the reception is logged as a '/'. This explains why the parameters for this efficacy score in the serve models are actually the second highest parameter amongst the efficacy parameters. After analyzing the corrected measure, no marginal differences were found from the results in table 7, therefore no changes are made to the presentation of table 7.

The negative and positive pressure point parameters are significant for the serve models, with the exception of female serves during a negative pressure point. Equal pressure point parameters are not found to be significant. Implications of these findings are discussed below table 7.

All thresholds are included in this model, due to the fact that all possible ordinal point indications are possible outcomes as a result of a serve, i.e. an ace (ordinal point indication = 6) and a fault serve (ordinal point indication = 1) are possible outcomes after a serve.

(ii) Reception models

The reception models have the lowest adjusted R^2 values, 0.366 for the male model and 0.486 for the female model. In contrast to the serve models, a logged '#' efficacy score for a reception does not imply a point is won due to this type of action. The '#' efficacy score parameter, set to 0, has therefore a smaller marginal difference towards the other parameters compared to the serve and attack models. This also explains the lower adjusted R^2 value.

The reception models show an increase in parameters' value for the efficacy scores according to table's 2 order, meaning that a logged reception as '=' is worse than '/' and a reception logged as '-' is better than a '/', and so forth.

The negative and positive pressure point parameters are significant for the reception models, with the exception of female receptions during a positive pressure point. Equal pressure point parameters aren't found to be significant. Implications of these findings are discussed below table 7.

A missing threshold in the reception models is threshold 5. Though a receiving team can win a point by a fault serve of the opponent, no reception can be logged for a fault serve. A serve goes out or in the net, for which in both cases no reception took place. Although a reception cannot be logged in case of an ace, a reception is logged when a serve is wrongfully received, in which case a direct point is scored on serve.[15] In that case, a reception is logged with a '=' efficacy score.

(iii) Attack models

The attack models have adjusted R^2 values of 0.580 for the male model and 0.549 for the female model. Similar to the serve models, a logged '#' efficacy score for an attack implies a point is won due to this action. The same is true for a '=' efficacy score, in which case a point is lost due to this action. The lower adjusted R^2 values for the attack models, compared to the serve models, is a result of the precise prediction of the ordinal point indications. A '=' or '#' serve efficacy score leads to a certain ordinal point indication of respectively '1' or '6', where within the attack models a '=' attack efficacy score leads to an ordinal point indication of '2' or '3' and a '#' attack efficacy score leads to an ordinal point indication of '4' or '5'.

The efficacy scores of an attack are accordingly to the described measures in table 2, with the exception of the '!' and '+' efficacies in the male model, where the parameter of the '!' efficacy score is actually higher (less negative) than the '+' efficacy score, while still significant. This

[15] Note that a 'direct point on serve' is indicated in table 3 as an ordinal point indication of '6', but for the receiving team, this ordinal scale is flipped, resulting in a ordinal point indication of 1, hence threshold 1 being applicable.

indicates that a male attack with an efficacy of '!' results in a higher probability of scoring a (high quality) point than a male attack with an efficacy of '+'.

The negative and positive pressure point parameters are significant for the attack models. Equal pressure point parameters aren't found to be significant. Implications of these findings are discussed below table 7.

In the attack models, the two outer thresholds are missing, i.e. threshold 1 and 5. Logically speaking, whenever an attack is logged it is not possible a point is indicated as an ace (or direct point on serve) or fault serve.

Table 7 - Ordered Probit models per action

	Male			Female		
	Serve	Reception	Attack	Serve	Reception	Attack
Negative pressure point	0.133*	0.240*	0.184*	-0.014	0.242*	0.138*
Positive pressure point	-0.231*	-0.146*	-0.265*	-0.245*	0.017	-0.146*
Equal pressure point	0.036	-0.039	-0.014	-0.059	0.067	0.018
Efficacy "="	-9.431*	-3.881*	-3.624*	-9.710*	-4.294*	-3.517*
Efficacy "/"	-2.574*	-1.023*	-3.236*	-3.710*	-0.757*	-3.163*
Efficacy "-"	-3.469*	-0.517*	-2.019*	-4.316*	-0.451*	-1.954*
Efficacy "!"	-3.297*	-0.358*	-1.766*	-4.222*	-0.171*	-1.874*
Efficacy "+"	-2.978*	-0.186*	-1.898*	-3.965*	-0.138*	-1.806*
Efficacy "#"	0[a]	0[a]	0[a]	0[a]	0[a]	0[a]
Threshold = 1	-6.497*	-2.580*		-7.073*	-2.639*	
Threshold = 2	-3.348*	-0.699*	-2.750*	-4.489*	-0.613*	-2.890*
Threshold = 3	-2.997*	-0.520*	-1.901*	-4.026*	-0.339*	-1.829*
Threshold = 4	-2.825*	-0.150*	-0.787*	-3.751*	0.145*	-0.534*
Threshold = 5	-1.251*			-1.692*		
N actions	20,591	16,740	23,621	13,814	12,155	17,912
Adjusted R^2	0.718	0.366	0.580	0.755	0.486	0.549

Notes: "a" This parameter is set to zero because it is redundant.
* Significant at the 1 percent level

I will elaborate the pressure point parameters below. In general, the efficacy parameters show high (less negative) values for high efficacies. So the better the performance, the higher the chance of scoring a (high quality) point. Exceptions or remarkable results regarding the efficacy parameters have been discussed within the according model elaborations above.

Implications of the pressure point parameters

In the male models, all positive and negative pressure point parameters are significant. In the female models the positive pressure point parameter is not significant in the reception model, whereas the negative pressure point parameter is not significant in the serve model. Both are closely related, where a serve on a positive pressure point is always followed – except for an ace or fault serve – by a reception on a negative pressure point. The equal pressure point parameters are not significant for any model in table 7.

The fact that all significant negative pressure point parameters have a positive value, and all significant positive pressure point parameters have a negative value shows that there is a difference in pressure between these pressure situations. During a negative pressure point, the probability of scoring a (higher quality) point is higher than when playing a positive pressure point. A positive pressure point appears to generate more pressure than a negative pressure point as a result of the so called ‘choking under pressure’, described in the theoretical section, resulting in a performance under par. This is in line with the theory of loss aversion. A lead in the in-game score results in a higher pressure situation for the leading team. Although on beforehand those (dis)utilities might seem fixed, whenever a team is close to a victory the reference to the disutility of losing a game results in that much additional pressure that performance drops. It seems that the disutility of losing a game is less evoked whenever a team is lagging in the in-game score, resulting in a difference in performance under negative and positive pressure points.

The fact that the equal pressure point parameters aren't significant in any model can be related to the fact both teams are under the same amount of pressure. Where the significant positive and negative pressure points indicate there is a difference in performance under these different pressure situations, for the equal pressure points, both teams relate to the same amount of pressure, or at least the same pressure creating situation. Therefore, the added pressure can be high, but because both teams encounter the same amount of pressure, both performances are under (or for that matter over) par and no significant difference is found between equal pressure points and no pressure points.[16]

Another remarkable result, is the marginal difference in magnitude of the significant positive and negative pressure point parameters within the male models. The negative pressure point parameter of the reception model (0.240) deviates more from the base variable – that is, no pressure point, set to 0 – than the positive pressure point parameter in the same model (-0.146). The opposite is true for the serve and attack male models. Where the reception and serve are closely related to one another – a good serve is a bad reception and vice versa – it appears, that when a player is under any form of additional pressure, a serve is more heavily underperformed than a reception. The same is true for an attack, under positive pressure the performance is marginally worse than the better performance under negative pressure.

[16] In a no pressure point situation both performances are expected to be on par.

Team's season performance

Given the significant Ordered Probit model fitting the volleyball game, I included the teams as an interaction term with the pressure point dummy variables, (see tables 8 and 9) to see if there is a difference in performance in relation to pressure between teams. Again, male and female analysis are separated and the analysis focuses only on actions logged as receptions.[17]

The variables on the right in tables 8 and 9 are the same variables as in the model of table 7 and the parameters are very similar. A different base variable for the non-pressure point results, rather logical, in different parameter values. In the left all teams are included with their corresponding season points – in descending order. The pressure point dummy that indicates no pressure point was played, shows parameters that correct for the general efficacy score parameters, so that the season winning team (indeed) shows a better significant base performance than others. In both the male and female model, note that many parameters of the team and pressure point interactions are significant. This is mainly the result of the smaller subgroups as a consequence of adding multiple teams to the equation. This makes a clean and clear comparison between closely equal performing teams difficult. In general, most negative pressure point parameters remain positive and most positive pressure point parameters remain negative.

Table 8 does show that the top four teams significantly perform better when not playing under pressure, which as a result, puts them at the top of the table. This confirms the face validity of my model. Only the runner up amongst the top four shows significant differences in performance when playing positive and negative pressure points.

[17] The reception is perceived to be the best action to measure in combination with team performance under pressure, where the efficacy scores are the least directly related to a score but expected to resonate most in the probability of scoring the next attack.

Table 9 especially shows that the bottom four women teams in the standing perform significantly worse in regular circumstances – the non-pressure points. Due to the smaller sample size of the women's data, less parameters are significant than in the male model.

Table 8 - Ordered Probit model – Reception - Male

Team - [points end of season]	Pressure point situations				Original variables	
	Negative	Equal	Positive	No		
Landstede Volleybal - [63]	0.087	-0.054	-0.154	0.115**	Threshold = 1	-2.591*
Abiant Lycurgus - [58]	0.381**	0.185	-0.385*	0.091***	Threshold = 2	-0.704*
Draisma Dynamo - [49]	0.054	0.013	0.057	0.092**	Threshold = 3	-0.524*
Orion - [42]	0.179	-0.350	-0.140	0.081***	Threshold = 4	-0.152*
Kootfin Taurus - [32]	0.145	0.043	-0.374***	-0.008		
Prins-VCV - [31]	0.500*	0.175	-0.173	0[a]	Efficacy "="	-3.890*
TT Papendal-Arnhem - [27]	0.099	0.123	-0.018	-0.073	Efficacy "/"	-1.026*
Inter Rijswijk - [25]	0.272***	-0.163	-0.072	-0.074	Efficacy "-"	-0.523*
ARBO Rotterdam-Fusion - [19]	0.270**	-0.303	-0.038	-0.006	Efficacy "!"	-0.359*
Rivo Rijssen - [18]	0.288***	-0.039	-0.079	-0.110**	Efficacy "+"	-0.183*
Zaanstad - [16]	0.346**	-0.433***	-0.111	-0.038	Efficacy "#"	0a
SkyLift-SSS - [16]	0.160	-0.350	-0.331***	-0.036		
					Adjusted R^2	0.369

Table 9 - Ordered Probit model – Reception – Female

Team - [points end of season]	Pressure point situations				Original variables	
	Negative	Equal	Positive	No		
Sliedrecht Sport - [43]	0.038	-0.044	-0.107	0[a]	Threshold = 1	-2.735*
VC Sneek - [39]	0.141	-0.057	0.039	-0.003	Threshold = 2	-0.704*
Coolen Alterno - [38]	0.353	0.179	-0.047	-0.068	Threshold = 3	-0.428*
Peelpush - [37]	0.055	-0.039	0.105	-0.078	Threshold = 4	0.059
Springendal Set-Up `65 - [36]	0.530**	0.129	-0.205	-0.060		
Eurosped TVT - [36]	0.227	0.237	0.127	-0.010	Efficacy "="	-4.291*
CSE VC Zwolle - [19]	0.456*	-0.300	-0.261	-0.101**	Efficacy "/"	-0.740*
TT Papendal-Arnhem - [11]	0.271	-0.788**	-0.119	-0.178*	Efficacy "-"	-0.449*
New Nexus Apps-Lycurgus - [8]	-0.163	0.232	-0.298	-0.192*	Efficacy "!"	-0.173*
Irmato-VC Weert - [3]	-0.562**	-1.026	-0.707***	-0.197*	Efficacy "+"	-0.137*
					Efficacy "#"	0[a]
					Adjusted R^2	0.489

Notes: "a" This parameter is set to zero because it is redundant.
* Significant at the 1 percent level
** Significant at the 5 percent level
*** Significant at the 10 percent level

V. Conclusion

General discussion

The negative and positive pressure point dummies being significant, makes it possible to discuss the research questions stated in the theoretical section, being: (i) does pressure influence a team’s performance and if so (ii) how do the different pressure points (negative, positive or equal) influence performance?

The first research question can be answered by pointing towards the many significant positive and negative pressure point parameters. The difference in pressure is the only variable in an otherwise same setting that can influence performance. Due to the fact that equal pressure points are played with both teams under the same amount of pressure, performance is logically not significantly different from the non-pressure situation in which case both teams perform equally well under the same amount of pressure, though little pressure. In other words, by the difference in performance between positive and negative pressure points we can extrapolate that pressure influences a team's performance. The parameters that do not support this statement are the female negative pressure point parameter and the opposing reception parameter on a positive pressure point due to their lack of significance.

This is in line with the theoretical section in which pressure influences performance along the likes of an inverted U shape in which the term ‘chocking under pressure’ is frequently mentioned as the reason of being the downfall in performance.

The second research question, on how different pressure points would influence performance, is answered by looking at the direction of the significant parameters as well as the structure of the thresholds. As mentioned in the results, the significant negative pressure point parameters show a positive value and the significant positive pressure point parameters show a negative value. The thresholds corresponding with the lower quality points, show lower threshold parameter values (more negative). The better the quality of a point, the higher a threshold parameter value is (less negative). The combination of the values of the negative and positive pressure point parameters and the threshold parameters results in a higher probability for a better performance when playing a negative pressure point. So, in general, a negative pressure point is more often won than a positive pressure point, ceteris paribus.

In the theoretical section, loss aversion theory is regarded. I argued that the team that is strongest influenced by loss aversion due to the in-game score, will feel the most pressure and is more likely expected to ‘choke under pressure’. It turns out, that the team in front (playing a positive pressure point) is most affected by loss aversion, which means that a team that is slightly in front in-game score at the end of a set experiences more pressure. Simply put, the team ahead feels the need to win it, now that they are ahead and close to victory. As a result, the game (or set) feels like it is theirs, in which case it can be moreover lost than won. Loss aversion theory states that the disutility of losing is absolutely bigger than the utility of winning, which shows to be influencing the performance under pressure in volleyball. In volleyball, a performance is always based on an interaction between the teams, so that a performances under or over par are always relative to one another.

Academic contribution

In current literature, few natural situations have been reviewed regarding pressure versus performance. Most literature uses mathematical based performance in an artificial pressure setting. The contribution of this paper firstly provides supporting research on this topic by enforcing high external validity with the use of data from a top volleyball competition. Secondly, few research efforts regarded loss aversion as a theory that influences pressure creating situations, here it does.

Suiting the external validity conditions, this paper supports current theory, that additional pressure can reduce performance. It is though fully related to the loss aversion theory where the performance of the two teams under pressure is always related to one another. In other words, the only factor that changed is pressure, though it changed for both teams. I theorize with the loss aversion theory that the leading team experiences the most additional pressure and therefore performs worse than par and worse than the opposition.

In this paper, and in relation to volleyball, the losing disutility described by the loss aversion theory is most evoked when a match (or set) is to be won, but there is still a reasonable chance for the opponent to win it. Literature that reviewed performance under pressure did not take loss aversion into account or did not find any reason to include it as an influencing factor. (See Cao, Price, & Stone (2011)) The presence of an opposing team on the other side of the pressure spectrum, i.e. they feel the match (set) is lost and can only win it from here on, induces a bigger contrast than yet has been reviewed. I argued that many situations in life occur with a form of opposition and contradicting pressure situations, in which I showed that loss aversion plays a substantial role.

In short, when it is yours to win, excel or perform, the pressure is higher in line with the loss aversion theory. When you have nothing to lose (because you are deemed to lose) pressure is less.

Managerial implications

Managerial implications are two folded. The first is to acknowledge pressure moments and interfere in consequence thereof. The second is to not create (unnecessary) pressure moments. Elaboration on these implications is not the scope of this research though, so that all implications are to be read as guidelines.

Acknowledgement of pressure moments and being wary of the related effects, ensure corresponding interference is adequate and well placed. The goal should be to reduce the pressure felt by a team, individual or perhaps yourself. I theorized that the prospect of a losing disutility influences performance for the worse and it is that disutility that should be managed accordingly. This paper showed that this is the case whenever a team is in front during a pressure point (i.e. chocking under pressure). The aim of this paper was not to review best ways of interference. Then again, it is believed that there is not a single solution for every team or individual. Some might prefer a direct approach by mentioning the situation and related pressure while others find more use in direct instructions enforcing them to focus on a single matter, thus not the pressure. When related to volleyball, time-outs are a useful tool to break the momentum and take away the pressure that might be felt. Again, this was not the scope of this paper.

A second managerial implication is not to create these moments when there is no need for. When delivering a product or service, as a team manager, this research shows that creating an atmosphere around "We have got nothing to lose" leads to better performances than "We can only mess it up". Although it is to be noted that for low pressure tasks added pressure can still increase performance. Summarizing, a general approach for how to deal with pressure is not available. It is a manager´s responsibility to acknowledge how different teams or individuals react and handle pressure to act accordingly. Being aware, applying methods such as mentoring and coaching, individuals as well as teams, arranging focus and alignment sessions, organizing timely, progress and evaluation meetings will help to identify and manage pressure points adequately and in time.

Limitations and directions for future research

Constructing quite some variables based on logical functions may have introduced some need for further and enhanced study. The main variable I needed to construct from other data because it was not present in the dataset, was the in-game score. From this in-game score I extracted the different pressure points. The noticeable errors I corrected already manually, which leaves the less noticeable errors that may have been present during the analysis. In future research a logged in-game score can easily eliminate this flaw.

Another limitation of this research was that multiple people logged the data. Although clear agreements were made on how to value different actions, there is always some room for individual interpretation. This comment applies to the efficacy scores included for each single action.

Within the volleyball game set-ups were eliminated, which can be perceived as a limitation. The ball touch between a reception and an attack is not measured on its quality by an efficacy score. The guiding principle is that a set-up is correct when a reception is correct, and poor when a reception is poor. In this paper a correct reception followed by a poor set-up, leading to a poor attack is interpreted as a failing performance of the attacker who should do well on a correct reception, though in reality it is the setter that delivered a poor performance. Vice versa the same is true and not controlled for in this paper.

With the current dataset some opportunities for future research can be identified. Although the quality of setters is not incorporated in the database, it is known what decisions setters make. This paper regarded an overall performance with cognition (a well overthought action) and physicality (physical skill e.g. power or technique) as the major cornerstones for performance without making a distinction between the two. The choices of a setter are mainly cognitive based decisions. Future research might focus on these decisions in relation to pressure moments.

Questions like; 'Do setters make different choices when under pressure?' can be an interesting addition to this paper.

Another focus might be on individuals and their performance under pressure. In this paper I only regarded total performance and performance of teams. If this dataset would allow it for individuals to retrieve enough data to get significant results it might be interesting to see how individuals with their demographic or volleyball specific properties handle pressure. For managerial implications it can be very useful to know what specific players never drop the ball when under pressure and what players do.

In addition to the given direction for future research, a different but interesting finding by Beilock & Carr (2005) is that individuals with a high working memory capacity are more (negative) influenced by pressure and anxiety than individuals with a low working memory capacity. They hypothesize that: *"If pressure and anxiety target individuals high in working memory capacity, this would carry significant implications for interpreting performance in high-pressure situations."* Memory capacity might be interesting to consider for possible future research in relation to physical performance and loss aversion theory.

References

Amabile, T. (1993). Motivational Synergy - Toward new Conceptiolizations of Intrinsic and Extrinsic Motivation in the Workplace. *Human Resource Management Review*, 185-201.

Apesteguia, J., & Palcios-Huerta, I. (2010). Psychological Pressure in Competitive Environments: Evidence from a Randomized Natural Experiment. *American Economic Review*, 2548-2564.

Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. (2009). Large Stakes and Big Mistakes. *The Review of Economic Studies*, 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*, 610-620.

Beilock, S. L., & Carr, T. H. (2005). When High-Powered People Fail: Working Memory and "Chocking Under Pressure" in Math. *Psychological Science*, 101-105.

Beilock, S., Kulp, C., Holt, L., & Carr, T. (2004). More on the fragility of performance: Choking under pressure in mathematical problem solving. *Journal of Experimental Psychology*, 584-600.

Camerer, C., Babcock, L., Loewenstein, G., & Thaler, R. (1997). Labor Supply of New York City Cab Drivers: One day at a Time. *The quarterly Journal of Economics*, 407-441.

Cao, Z., Price, J., & Stone, D. (2011). Performance Under Pressure in the NBA. *Journal of Sports Economics*, 231-252.

Francesco, F., Innocenti, A., & Pin, P. (2013). Is there psychological pressure in competitive environments? *Journal of Economic Psychology*, 249-256.

Hardy, L., & Parfitt, G. (1991). A catastrophe model of anxiety and performance. *Britisch Journal of Psychology*, 163-178.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 263-292.

Kocher, M., Lenz, M. V., & Sutter, M. (2012). Psychological pressure in competitive environments: New evidence from randomized natural experiments. *Management Science*, 1585-1591.

Markman, A. B., Todd Maddox, W., & Worthy, D. A. (2006). Choking and Excelling Under. *Psychological Science*, 944-948.

Mas, A. (2006). Pay, Reference Points, and Police Performance. *Quarterly Journal of Economics*, 783-821.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 459-482.