Psychology of Sport and Exercise 14 (2013) 531–537

Contents lists available at SciVerse ScienceDirect

Psychology of Sport and Exercise

journal homepage: www.elsevier.com/locate/psychsport

Pitching and clutch hitting in Major League Baseball: What 109 years of statistics reveal

Mark P. Otten*, Matthew E. Barrett

California State University, Northridge, USA

A R T I C L E I N F O

Article history:
Received 30 August 2012
Received in revised form
1 March 2013
Accepted 1 March 2013
Available online 14 March 2013

Keywords:
Clutch performance
Anxiety
Motor skills
Baseball
History
Archival study

A B S T R A C T

Objectives: Theory on performance under pressure in sport has proposed that an athlete may be disrupted psychologically when distracted, or when explicitly monitoring too much the skills involved (Beilock & Carr, 2001; Masters, 1992). Research has also suggested that the extent to which an athlete allows pressure to impact performance may be greater for skills of increased complexity, such as hitting a baseball (Kinrade, Jackson, Ashford, & Bishop, 2010; Masters, Polman, & Hammond, 1993). Accordingly, hypotheses for the current study were that baseball hitters would be more susceptible to pressure-induced performance changes than pitchers, whose skills are less based in hand-eye coordination.
Design & method: An archival design was employed, accounting for 109 years of historical baseball data at both the team and individual levels.
Results: In line with hypotheses, for players with a minimum of 10 postseason innings pitched in a single year ($n = 835$) pitching statistics were significantly correlated from regular season (less pressure) to postseason (more pressure). For those with a minimum of 20 postseason at bats in a year ($n = 1731$), hitting statistics were similarly correlated from season to postseason; overall, however, the weakest such relationship was batting average. For teams ($n = 370$), regular season pitching was expected to be the best predictor of postseason success rates; this hypothesis was supported, but only for the most recent era of baseball history (1995–2011).
Conclusions: The data imply that, while hitting should not be wholly neglected, a successful, clutch baseball team should be built primarily around pitching.

Much has been made of the effects of elevated pressure in sport, from its potential to induce an athlete to “choke” (Baumeister, 1984; Gray, 2004; Jackson, Ashford, & Norsworthy, 2006; Mesagno, Harvey, & Janelle, 2011) to its prospective positive impact if interpreted adaptively (Cheng, Hardy, & Markland, 2009; Otten, 2009). Indeed, an athlete’s performance is often disrupted, one way or another, under high-stakes conditions. Much theory has been proposed to explain this disruption and its psychological foundations. Distraction theory (DeCaro, Thomas, Albert, & Beilock, 2011; Wine, 1971) proposes that pressure deflects an athlete’s attention away from the task at hand, causing performance to suffer. Alternatively, explicit monitoring theory (Beilock & Carr, 2001; Masters, 1992) posits that conscious attention paid to the step-by-step processes of a motor skill may trigger a performance letdown under pressure.

Baseball’s pressure-filled World Series dates back to 1903 in the United States, when the Boston Americans (now known as the Red Sox) of the American League defeated the Pittsburgh Pirates of the National League, five games to three. The event established a rivalry between the leagues, and played to overflow crowds. After a one-year hiatus, the Series was set as an annual event starting in 1905, to establish the champion of Major League Baseball. The teams with the best records in each league qualified for the event, which stood as baseball’s only postseason series until 1969. From 1969 until 1993, additional League Championship Series were added, and in 1995 (after a player strike halted play in 1994) two Division Series per league were further included to serve as an additional round of playoffs.

Thus, with the championship on the line, we may collect data from such postseason play to stand for baseball’s version of performance under heightened pressure conditions. Indeed, most traditional forms of pressure tend to rise in the postseason (e.g., fan interest and attendance, media attention, internal and external awards; see Baumeister & Showers, 1986). The regular season play that leads up to these playoffs, then, might be considered low-pressure (or at least, non-elevated pressure) conditions. As such, we can compare the two and derive predictions for which players

* Corresponding author. 18111 Nordhoff St., Northridge, CA 91330-8255, USA. Tel.: +1 310 922 2734.
E-mail address: mpotten@csun.edu (M.P. Otten).

1469-0292/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.psychsport.2013.03.003

(and teams) might be more or less successful under the bright lights of the postseason, relative to their own same-year performances during the regular season. Baumeister's (1984) definition of a choke implied a performance decrement that occurs under high anxiety, relative to one's own standards. Likewise, Otten (2009) proposed that a "clutch" performance is a similar performance increment under pressure conditions.

As Kinrade et al. (2010) note, much of the existing literature on choking has been derived from empirical studies of relatively complex motor skill tasks that require considerable technical instruction (e.g., Beilock & Carr, 2001; Jackson et al., 2006). Included among these studies is that of Gray (2004), who found that poorer performance under pressure among a sample of skilled hitters in baseball was associated with greater explicit monitoring of the position of the bat. Indeed, hitting a baseball at the highest level is often considered by experts as the most difficult, complex skill in all of sports (see Peterson, 2011). Thus, that popular theories of pressure and performance are applicable to hitting a baseball in postseason play is considered a safe assumption here.

For pitching a baseball, however, this assumption may not be quite so clear. While it is by no means a simple skill, pitching likely requires a different mental approach, involving greater calculated effort, perhaps, than the split-second hand-eye coordination required for hitting. Thus, it could be that explicit monitoring of one's pitching motion mechanics could be *less* detrimental to a pitcher's performance under pressure than such monitoring of skills by a hitter, as in Gray (2004). This is not to say that pitchers feel less pressure, or respond to it with a reduced tendency to consciously monitor their movement. Rather, given the same amount of explicit monitoring of skills, hitters' performance may be impacted more due to the complexity involved in reacting to the ball thrown at high speed. To date, no research has examined these relationships *within* a specific sport such as baseball.

However, it is conceivable that the contrast between the complexities involved in pitching versus hitting may approximate the differences observed by Kinrade et al. (2010) and Masters et al. (1993) in their comparisons of motor skill and cognitive tasks. Masters et al. (1993) found that reinvesting attention, or more actively attending to the task, was associated with performance decrements under pressure among separate samples of golf putters, tennis and squash players. The same effect was *not* found, however, among participants asked to complete a two-dimensional rod-tracing task, leading Masters et al. (1993) to conclude that explicit monitoring was not such a bad thing for those executing a simpler task. In support of these findings, results from Kinrade et al. (2010) suggest that greater reinvestment of attention to a cognitive task may speed up one's performance, but not necessarily degrade it under pressure.

Taken together, the above findings lead us to believe that *not* explicitly monitoring your skills while under pressure may be more important for hitters than for pitchers in baseball. If this is true, then other psychological variables central to explicit monitoring theory (e.g., feelings of perceived control; Cheng et al., 2009; Otten, 2009) should also be more relevant for hitters. Thus, the successful hitter would be one who is best-prepared psychologically, regardless of say, his level of success achieved during the regular season.

The current study

At the individual level, pitching skills will be consistent across the regular season, playoffs and World Series. That is, individual regular season pitching statistics will be significantly correlated with the same postseason pitching statistics, across pitchers who participated in postseason play. For hitters, however, skills will be less consistent, and thus season–postseason correlations will be significantly smaller than those for pitching (but still, significantly different than zero). One exception should be noted: the stolen bases category is also listed among hitting statistics (see below). Since base-stealing is *not* part of the act of hitting and is less reliant on hand-eye coordination, much like pitching, it is considered a less complex skill here and is thus hypothesized to hold a stronger season–postseason correlation than the other hitting stats.

The logic of the above hypotheses may be extended to the team level as well. If pitching is relatively more stable across the regular year and postseason, it follows that teams that collectively pitched well during the season will not only also pitch well under pressure, but also *win postseason games*. As a result, regular season team hitting will be significantly less predictive of postseason team success than will regular season team pitching.

We also note that there is an old adage in baseball that pitching and defense win championships (e.g., Merron & Schoenfield, 2005). The above hypotheses are consistent with the old adage, in that pitching skills are proposed as significantly stronger predictors of championship success than hitting skills.

Method

Design & materials

To test these hypotheses, an archival research paradigm was employed. The data were downloaded from the Baseball Reference website (Baseball-Reference.com, 2012), where a plethora of baseball data is available for public access and use. The data retrieved span 109 years of baseball history, from the inception of the World Series in 1903 through to the 2011 playoff series. All individual cases included in the data set were professional baseball players who collected a minimum of either 20 postseason at bats, for hitters ($n = 1731$) or 10 postseason innings pitched, for pitchers ($n = 835$) in a given year. The data were paired so as to list regular season statistics alongside postseason stats, by year, for each player included. A separate data set was comprised of team-level statistics, encompassing team hitting and pitching for every team in history that had played at least one postseason series ($n = 370$). Data were coded and screened using Microsoft Excel.

As noted earlier, prior to 1969 only the World Series was contested, while additional rounds of playoffs have been added in recent years; as a result, some players and teams in the sample played in multiple postseason series in a single year. Thus, totals were summed (e.g., for home runs) and/or averaged (e.g., for batting average) to achieve only one listing per player or team, per year. For example, Josh Hamilton appeared in three postseason series per year in 2010 and 2011, but he is only listed twice in the final data set, once for 2010 and once for 2011.

Measurement

Hitting statistics. Total *at bats* tallies a hitter's total plate appearances, minus select cases (e.g., when he walks or is hit by a pitch). The *hits* statistic counts the total number of times a batter puts the ball in play and reaches first base, not due to a fielder's choice or error. A *home run* occurs when a player scores based on his own hit, and not due to a fielding error. While running the bases, a player may achieve a *stolen base* when he advances while the pitcher is delivering the ball. *Batting average* is the percentage of successful hits accumulated per at bat; *on base percentage* is the percent of time a player reaches base successfully per at bat, as it is possible to reach base without a hit (e.g., by a walk). The *slugging percentage* for a player is the total number of bases reached divided by at bats; for example, a single counts as one base, while a home run counts as four. Finally, the *on base plus slugging percentage (OPS)* is the sum of these two statistics, and is often regarded as an indicator of the overall offensive ability of a hitter.

Pitching statistics. For pitchers, the *innings pitched* statistic counts one's total number of innings completed. The *hits* statistic reflects the number of hits given up to the opposition; walks, or *bases on balls* tallies those to whom a pitcher yields four balls. *Earned runs* is the number of runs given up by a pitcher, not due to fielding errors. The *earned run average (ERA)* is a popular indicator of overall pitching ability, and is the number of earned runs, multiplied by nine, divided by innings pitched. Also indicative of overall pitching aptitude is the *walks plus hits per inning pitched (WHIP)* statistic, which is calculated by the total number of walks and hits divided by innings pitched.

Team-level statistics. The aforementioned statistics for hitting and pitching were also contained in the team-level data set, along with an additional few. *Wins* and *losses* were included for teams, and a *winning percentage* was calculated from these for both regular season and postseason. The latter was then used as a dependent variable representing team postseason success.

To balance sample sizes, several statistics were also coded on a per-at bat or per-9 innings basis for individuals, and a per-game basis for teams. Data were then analyzed using SPSS version 19.

Results

Descriptive statistics: summary

Tables 1 and 2 present descriptive statistics for individual pitchers and hitters, respectively. To achieve a ranking of "clutch" performance, indices of success were chosen for pitchers (WHIP) and hitters (OPS); difference scores were then calculated for each player, subtracting his regular season value from his postseason value. For example, the Yankees' Allie Reynolds accumulated a 1.51 value in the 1949 regular season and a .49 value in the postseason, for a difference before rounding of −1.03 (.49 − 1.51). Note that WHIP was chosen over ERA here due to a potential floor effect for the latter statistic, since 35 of the 835 pitchers sampled achieved postseason ERA values of .00. The best and worst five pitchers (Table 1) and hitters (Table 2) are then displayed according to these ranks. There were no missing data.

Descriptive statistics: individual pitchers

Overall, individual pitchers in our sample averaged 17.61 innings pitched ($SD = 6.72$) in the postseason, with a mean ERA of 3.14 ($SD = 1.96$) and WHIP of 1.20 ($SD = .34$); in the regular season, they averaged 203.12 innings pitched ($SD = 68.60$), an ERA of 3.16 ($SD = .84$) and a WHIP of 1.21 ($SD = .15$). A repeated-measures MANOVA with ERA and WHIP serving together as dependent variables suggested no significant mean difference in pitching success across regular season to postseason, $F(1, 831) = .24$, $MSE = 1.17$, $p = .63$.[1]

Table 1
Descriptives: five most and least "clutch" individual pitchers of all time (minimum 10 postseason innings pitched (IP); $n = 835$).

Year	Name	Regular season			Postseason			WHIP-diff.
		IP	ERA	WHIP	IP	ERA	WHIP	
Top five								
1949	Allie Reynolds[1]	213.2	4.00	1.51	12.1	0.00	0.49	−1.03
2004	Derek Lowe[2]	182.2	5.42	1.62	19.1	1.86	0.72	−0.89
1996	David Weathers[1]	88.2	5.48	1.69	11.0	0.82	0.82	−0.87
1956	Don Larsen[1]	179.2	3.26	1.28	10.2	0.00	0.47	−0.81
1947	Hugh Casey[3]	76.2	3.99	1.36	10.1	0.87	0.58	−0.78
Bottom five								
1993	Tommy Greene[4]	200.0	3.42	1.19	11.2	13.11	2.57	+1.39
2007	CC Sabathia[5]	241.0	3.21	1.14	15.1	8.80	2.22	+1.08
1980	Marty Bystrom[4]	36.0	1.50	0.97	10.1	3.48	1.94	+0.96
1985	Joaquin Andujar[6]	269.2	3.40	1.29	14.1	7.53	2.23	+0.95
2004	Javier Vazquez[1]	198.0	4.91	1.29	11.1	9.53	2.21	+0.92

Note. [1]New York Yankees; [2]Boston Red Sox; [3]Brooklyn Dodgers; [4]Philadelphia Phillies; [5]Cleveland Indians; [6]St. Louis Cardinals.

Table 2
Descriptives: five most and least "clutch" individual hitters of all time (minimum 20 postseason at bats (AB); $n = 1731$).

Year	Name	Regular season			Postseason			OPS-diff.
		AB	BA	OPS	AB	BA	OPS	
Top five								
2008	Manny Ramirez[1]	552	.332	1.031	25	.520	2.040	+1.009
1989	Rickey Henderson[2]	541	.274	.810	34	.441	1.647	+.837
1990	Billy Hatcher[3]	504	.276	.708	27	.519	1.481	+.773
1953	Billy Martin[4]	587	.257	.710	24	.500	1.478	+.768
2004	Carlos Beltran[5]	599	.267	.915	46	.435	1.652	+.737
Bottom five								
1911	Red Murray[6]	488	.291	.781	21	.000	.087	−.694
1952	Gil Hodges[7]	508	.254	.886	21	.000	.192	−.694
1962	Mickey Mantle[4]	377	.321	1.091	25	.120	.401	−.690
2003	Eric Chavez[8]	588	.282	.864	22	.045	.178	−.686
1934	Marv Owen[9]	565	.317	.837	29	.069	.169	−.668

Note. [1]Los Angeles Dodgers; [2]Oakland Athletics; [3]Cincinnati Reds; [4]New York Yankees; [5]Houston Astros; [6]New York Giants; [7]Brooklyn Dodgers; [8]Oakland Athletics; [9]Detroit Tigers; BA = batting average.

Descriptive statistics: individual hitters

The individual hitters averaged 33.56 postseason at bats ($SD = 12.40$), with a mean batting average of .258 ($SD = .079$) and OPS of .742 ($SD = .243$); in the regular season, 483.74 was the mean number of at bats ($SD = 119.36$) in the sample, with a mean batting average of .282 ($SD = .031$) and OPS of .791 ($SD = .111$). Here, a repeated-measures MANOVA with batting average and OPS serving as dependent variables did suggest a significant mean decrease in hitting success from regular season to postseason, $F(1, 1730) = 100.08$, $MSE = .02$, $p < .001$; partial $\eta^2 = .06$.[2]

Descriptive statistics: teams

Table 3 shows descriptive stats for teams ranked according to winning percentage difference scores. That is, a team's regular season rate of success is subtracted from its postseason clip, for each year. For example, the 1922 New York Giants' regular season winning percentage was .604, but it was 1.000 in the postseason, since they swept the Yankees in the World Series that year four games to none. Thus, the '22 Giants' difference score was .396 (1.000 − .604). The best and worst teams are then displayed in this order in the table. On average, teams in the sample averaged 3.59 wins ($SD = 2.91$) and 3.60 losses ($SD = 1.53$) in the postseason, and a winning percentage of .434 ($SD = .257$); in the regular season, teams won an average of 94.77 games ($SD = 7.93$) while losing only 62.50 ($SD = 7.99$), with a winning percentage of .603 ($SD = .042$). A dependent-samples t test to compare these two winning percentages suggested a significant decrease from regular season to postseason, $t(369) = 13.18$, $p < .001$.

[1] Separate repeated-measures ANOVAs revealed no mean difference in ERA [$F(1, 834) = .09$, $MSE = 1.81$, $p = .77$] nor WHIP [$F(1, 831) = 1.44$, $MSE = .06$, $p = .23$] across regular season to postseason.

[2] Separate repeated-measures ANOVAs revealed significant mean differences in batting average [$F(1, 1730) = 156.08$, $MSE = .00$, $p < .001$] and OPS [$F(1, 1730) = 77.79$, $MSE = .03$, $p < .001$] across regular season to postseason.

Table 3
Descriptives: five most and least "clutch" teams of all time ($n = 370$).

Year	Name	Regular season			Postseason			Pct.-diff.
		W	L	Pct.	W	L	Pct.	
Top five								
1922	New York Giants	93	61	.604	4	0	1.000	+.396
1966	Baltimore Orioles	97	63	.606	4	0	1.000	+.394
1963	Los Angeles Dodgers	99	63	.611	4	0	1.000	+.389
1914	Boston Braves	94	59	.614	4	0	1.000	+.386
1976	Cincinnati Reds	102	60	.630	7	0	1.000	+.370
Bottom five								
1954	Cleveland Indians	111	43	.721	0	4	.000	−.721
1914	Philadelphia Athletics	99	53	.651	0	4	.000	−.651
1963	New York Yankees	104	57	.646	0	4	.000	−.646
1980	New York Yankees	103	59	.636	0	3	.000	−.636
1939	Cincinnati Reds	97	57	.630	0	4	.000	−.630

Note. W = wins; L = losses; Pct. = winning percentage.

Linear analyses: individual players

It was hypothesized that at the individual level, pitching skills would be more consistent across regular season and postseason than would hitting skills. To address this conjecture, presented in Table 4 are regular season versus postseason bivariate correlations across all individual pitchers and hitters sampled. For example, postseason ERA correlated with regular year ERA at $r = .284$ ($p < .001$) for all $n = 835$ pitchers in our sample. Tests of differences between these correlations (see Lowry, 2012) revealed that the regular season–postseason correlation for batting average was significantly lower than all other correlations shown in Table 4, except for one (WHIP; $p < .05$). Likewise, the WHIP correlation was significantly less than the highest two (stolen bases per at bat and OPS; $p < .05$); the ERA and OPS correlations were also significantly lower than that of stolen bases per at bat ($p < .05$).

To further assess individual-level hypotheses, Table 5 supplements these analyses to correlate individual regular season statistics with regular season–postseason difference scores (as seen in Tables 1 and 2). For example, the higher a pitcher's regular season ERA was, the more his ERA dropped from season to postseason ($r = -.15$, $p < .001$).

Linear analyses: teams

At the team level, it had been hypothesized that teams that collectively pitched well during the regular season would go on to enjoy the greatest success in the postseason. To test this projection, a series of team-level multiple linear regressions were conducted, with regular season ERA (to represent pitching success), OPS (to represent hitting success) and stolen bases per game set as predictors of postseason winning percentage. One such multiple regression was performed for each of four time periods across the history of baseball, as shown in Table 6. Each variable was screened for normality, and ERA, OPS and stolen bases per game were cleared of multicollinearity before entry into each regression equation. To follow up these analyses, bivariate correlations were run for each time period, to involve additional regular season team-level statistics alongside postseason winning percentage. A map of these correlations is shown in Table 7.

Table 4
Regular season–postseason correlations.

Individual pitchers (n = 835)	
ERA	.284***
WHIP	.242***
Individual hitters (n = 1731)	
Stolen bases per at bat	.575***
Batting average	.172***
OPS	.326***

Note. ***$p < .001$.

Table 5
Regular season–postseason difference score correlations.

Individual pitchers (n = 835)	
ERA	−.149***
WHIP	−.189***
Individual hitters (n = 1731)	
Stolen bases per at bat	−.161***
Batting average	−.221***
OPS	−.136***

Note. ***$p < .001$.

To supplement these team-level analyses, Table 8 shows postseason winning percentage and its correlations with regular season–postseason difference scores, split across historical time periods. For instance, across 1995–2011, the more a team's regular season batting average increased from season to postseason, the better that team did in the postseason ($r = .48$, $p < .001$).

Discussion

Individual-level hypotheses

For individual pitchers, we had hypothesized that statistics would be consistent across the regular season and postseason in a given year, since pitching represents a comparatively less complex skill that is therefore less susceptible to pressure-induced disruption (Beilock & Carr, 2001; Kinrade et al., 2010; Masters, 1992; Masters et al., 1993). Support was found for this hypothesis, in that all individual-level pitching statistics were significantly correlated across our sample. By the same theory, we had projected hitting to be less consistent from season to postseason, since it represented a more complicated task and was thus more vulnerable to psychology-related fluctuations under pressure. Support was *not* found directly for this hypothesis, in that all hitting statistics sampled were likewise significantly correlated from season to postseason. The one hypothesized exception had been stolen bases, since this represented a lower-complexity skill; this projection was supported, since the regular season–postseason correlation for this statistic was also significant.

Delving a bit further, we note that batting average did represent the lowest bivariate correlation among all individual-level statistics for our sample (see Table 4). Thus, while provisional by nature, this

Table 6
Team-level linear regressions to predict postseason winning percentage, by era.

IV: Team Statistic	*B*	*SE B*	*β*
[1]*1903–1918 (n = 30)*			
Regular season ERA	0.03	0.23	.04
Regular season OPS	−0.65	2.16	−.10
Regular season SB per game	−0.07	0.17	−.09
[2]*1919–1968 (n = 100)*			
Regular season ERA	−0.18	0.09	−.28
Regular season OPS	2.55	0.85	.44**
Regular season SB per game	0.98	0.13	.08
[3]*1969–1993 (n = 104)*			
Regular season ERA	−0.13	0.08	−.20
Regular season OPS	1.25	0.90	.17
Regular season SB per game	0.12	0.08	.16
[4]*1995–2011 (n = 136)*			
Regular season ERA	−0.23	0.06	−.41***
Regular season OPS	2.26	0.79	.30**
Regular season SB per game	−0.04	0.10	−.04

Notes. [1]$R^2 = .02$; [2]$R^2 = .09$; [3]$R^2 = .05$; [4]$R^2 = .10$; **$p < .01$; ***$p < .001$; IV = independent variable; SB = stolen bases.

Table 7
Team-level correlations with postseason winning percentage, by era.

Team Statistic	1903–1918 ($n = 30$)	1919–1968 ($n = 100$)	1969–1993 ($n = 104$)	1995–2011 ($n = 136$)
Reg. season SB per game[h]	−.133	.060	.147	−.020
Reg. season batting average[h]	−.217	.056	−.016	−.049
Reg. season OPS[h]	−.108	.231*	.037	.048
Reg. season ERA[p]	−.026	.056	−.111	−.215*
Reg. season WHIP[p]	−.071	.149	−.087	−.195*

Notes. [h]-hitting statistic; [p]-pitching statistic; *$p < .05$; SB = stolen bases.

suggests that the act of hitting a baseball in its simplest form (i.e., getting a hit of any kind) may indeed be more susceptible to performance changes due to pressure. It may be that hitting with power (e.g., earning a high OPS) serves as more of a specialty skill here (like stolen bases) and appears more consistent, since only a subset of hitters are capable of such achievement. We note also that when history is split into discrete time periods (as in Tables 6 and 7), batting average remains the lowest-ranked individual-level correlation for each era, with the exception of OPS during the so-called "dead-ball era" (1903–1918). The latter should be interpreted with extreme caution, however, since a potent floor effect for OPS seemed to be operating in these early years. Indeed, of the 121 hitters who accrued 20 or more postseason at bats prior to 1919, only 14 hit as many as one home run, and nobody hit more than two.

Note that minimums of 20 at bats and 10 innings pitched were set for postseason hitters and pitchers, respectively, to protect against spurious effects that may have been observed due to small sample sizes. For example, Brett Myers of the Phillies may have appeared extraordinarily clutch for his perfect 3-for-3 performance in the 2003 postseason, especially since his regular season average was a mere .069. However, his performance was much more likely spurious than say, that of Carlos Beltran in 2004, whose postseason output that year spanned 46 plate appearances (see Table 2). Still, some players with relevant data may have been excluded by these criteria (e.g., hitters with more than 10 at bats whose team was swept from the first round of postseason play). This represents a potential limitation to external validity for the current study; meanwhile, future researchers might add these data back in, and reanalyze for comparison.

We also note here that our main hypotheses test the relative stability of hitting and pitching performances across regular season and postseason. Meanwhile, absolute stability in means was found in pitching statistics (ERA, WHIP) as well; for hitting, however, batting average and OPS means dropped significantly from regular season to postseason. This suggests that on the whole, hitting became significantly more difficult under the pressure of the postseason, while pitching did not.

Table 8
Team-level difference score correlations with postseason winning percentage, by era.

Team Statistic	1903–1918 ($n = 30$)	1919–1968 ($n = 100$)	1969–1993 ($n = 104$)	1995–2011 ($n = 136$)
SB per game difference[h]	.243	−.048	−.107	.045
Batting average difference[h]	.659***	.387***	.495***	.484***
OPS difference[h]	.640***	.429***	.592***	.607***
ERA difference[p]	−.526**	−.606***	−.526***	−.411***
WHIP difference[p]	−.552**	−.541***	−.481***	−.360***

Notes. [h]-hitting statistic; [p]-pitching statistic; ***$p < .001$; **$p < .01$; SB = stolen bases.

Team-level hypotheses

On the team level, we had predicted that teams that collectively pitched well during the season would go on to enjoy postseason success. This hypothesis was supported, but only by the data from the most recent period of playoff baseball (1995–2011). For these years, regular season pitching success (as measured by ERA) was revealed as a significant predictor, in the expected direction, as part of a multiple regression depending on postseason winning percentage. Bivariate correlations also suggested one other index of regular season pitching success (WHIP) to be significantly related to postseason success. For years prior to 1995, however, regular season pitching statistics did not relate significantly to postseason success rates.

Our complementary team-level hypothesis was that regular season team hitting would serve as a comparatively weaker predictor of postseason team success. This projection was fully supported, again, for the years 1995–2011. For this period, regular season hitting success (as measured by OPS) significantly predicted postseason win percentage alongside successful regular season pitching (as indicated by ERA) as part of the aforementioned multiple regression; however, bivariate correlations revealed no significance for any regular season offensive stats with postseason success. This result suggests a form of classical suppression, whereby adding pitching to the equation as an independent variable caused the relationship between hitting and winning percentage to appear larger. In the process, variance that was irrelevant to the prediction of postseason success (e.g., that of other regular season statistics) was suppressed. In the end, we can say that season hitting success was related to postseason win percentage, but only when season pitching variance was already accounted for in the equation. Meanwhile, measured as a direct correlation, regular season hitting and postseason success rate were unrelated. This suggests that, based on recent years, successful postseason teams should be built around pitching first, and hitting second.

For periods from 1903 to 1918 and from 1969 to 1993, our regular season hitting-related hypotheses were also upheld, in that significant relationships with postseason success were not found. For our 1919–1968 sample, however, regular season OPS emerged as significantly correlated with postseason winning percentage. Note that time periods were stratified as such to account for changes to baseball's postseason format across time, as mentioned above. Due to these changes, playoff winning percentage actually decreased on average over time, since the total number of playoff teams increased while still only one champion was crowned. Thus, linear analyses necessarily accounted for these time trends, or else risked confounding by way of the winning percentage or other variable(s). As it turned out, several team-level statistics also trended up (OPS, ERA, WHIP) or down (stolen bases) across time; the numbers of teams and players featured in the league also increased over the years, thus providing us a number of reasons to believe analyses should be stratified.

Also, we note here that for the purposes of multiple regression analyses (see Table 6) only one pitching statistic (ERA) and one hitting statistic (OPS) were chosen to represent overall team pitching and hitting performance, respectively. This was done for fear of multicollinearity if multiple variables of the same type were chosen to serve as predictors in the same equation. For example, a home run simultaneously raises a hitter's batting average and OPS. Additional statistics were then re-included for bivariate correlation analyses (see Table 7). The stolen bases category was free of strong correlations with other predictors, and was thus included in the regression analyses; no prior hypotheses were made for it at the team level, however, and no significance was found.

Thus, we may conclude that team-level hypotheses were upheld for the years since the Division Series were added to baseball's playoffs in 1995. Prior to this period, however, support was mixed. With the additional games played, we might say that postseason sample sizes were larger in the later years, and thus conclusions are more reliable. Indeed, if additional playoffs had been present in the early days, would our conclusions have been different?

Certainly, changes in the way the game is played over time have been numerous, as noted above. Since leagues featured fewer total players and teams in years past, it may be that more players in recent days are of reduced expertise and, thus, less prepared to perform well under pressure. Then again, baseball has also become a more global game over time, and so the pool of experts for the leagues to draw from has also likely increased, perhaps canceling this hypothesis. Statistically, we know hitters have tended to hit for more power and steal fewer bases as time has gone by. Good pitching has become scarcer, and perhaps simultaneously more important as a predictor of postseason success.

The nature of pressure has likely also changed for baseball players and teams over time. Prior to 1950 or so, results were printed in newspapers and games were broadcast on radio only, if at all. While crowds were comparable to more recent standards, salaries were much more modest. With the advent of television, national and international media present and the multi-million-dollar business features of recent years, it is quite possible that players feel more (or at least different) pressure when on the field. By extension, we might suggest that the theory proposed here to suggest that hitters are most susceptible to pressure-induced performance changes may only apply when the level of pressure reaches recent standards, brought on by the unique anxieties produced by the modern game. The data seem to support this conjecture, though further research is needed to validate such claims.

Descriptive statistics: who is (or is not) clutch?

As mentioned earlier, definitions of choking or clutch performance tend toward the idea that under pressure an athlete performs better or worse than usual, relative to his or her own standards (Baumeister, 1984; Mesagno et al., 2011; Otten, 2009). Thus, for individual-level analyses we measured players here by postseason performance, relative to one's own regular season performance. As such, players may be sorted by regular season—postseason difference scores, as was done for Tables 1 and 2. By these ranks and current theory, we might expect the hitters with the highest ranks to be no different in their prior regular season success rates than those hitters with the lowest ranks, since hitting is more dependent on pressure than pitching.

Table 5 provides some further insight into these projections. For these analyses, we correlated the aforementioned difference scores with regular season statistics. All correlations here were in fact negative, suggesting perhaps regression to the mean, in that the more extreme a player's regular season value was, the greater the difference between it and his postseason value. Alternatively, we might interpret this as subpar regular season performers tending toward clutch performance more so than those who were more successful; this would be consistent with prior hypotheses for hitters, but not so for pitchers.

Descriptive statistics by rank for teams (see Table 3) yielded the best and worst clutch teams over time, according to regular season—postseason difference scores for winning percentage. The ends of this list lean toward teams from further in the past, since with fewer postseason series present, teams were more likely to go undefeated/winless and thus achieve extreme scores on one end or the other for percentage. Still, by original hypotheses one might expect teams that pitched well during the regular season to be more numerous toward the top of this list.

Table 8 provides some insight on this. Here, postseason winning percentage was correlated with regular season—postseason difference scores at the team level, separately for each historical time period. As is evident in the table, this relationship was significant for indices of hitting success (batting average and OPS) across all eras; it was also significant for indices of pitching success (ERA and WHIP), however, suggesting also that the more a team's pitching staff improved from season to postseason, the more success a team had.

What 109 years of statistics reveal: recommendations

Certain limitations inevitably exist with the current study, due to its archival nature. As such, key variables were not manipulated empirically, and thus conclusions must be noted as correlational and qualified accordingly. Ideally, we would have also acquired retroactive psychological data from the players and teams in question; we would have asked them about their feelings of control and tendencies to explicitly monitor their skills under pressure, and then correlated these measures with postseason success in order to clarify conclusions. Short of this, however, the current study is unique in its treatment of historical data across baseball history, and its application of pressure-performance theory in sport there.

Future research may explore the current data further by way of multilevel modeling. Players here were nested within teams, allowing for a hierarchical structure to the data with players set at level one, and teams set at level two. Analyses may then allow additional research hypotheses to be answered; for example, it could be proposed that individual regular season pitching success was a better predictor of postseason pitching success for some teams more than for others. Such a prediction crosses from the player level over to the team level. The current set of player-level hypotheses (e.g., regular season pitching success correlates with postseason pitching success) was handled entirely at the individual level, and was thus equivalent to a set of level-one hypotheses by this multilevel framework. Team-wide hypotheses were similarly contained at level two. Thus, with current hypotheses contained entirely within levels, separate linear analyses at both player and team levels were employed for simplicity, since within-level predictions under the multilevel model are theoretically equivalent. That said, it should be noted that team-level conclusions are not wholly independent of those at the individual level, as one group is nested within the other.

Throughout, we have set regular season and postseason play to operationally define situations of less and more pressure, respectively. This seems a reasonable assumption, since as suggested earlier, in postseason play most traditional forms of pressure (Baumeister & Showers, 1986) seem to rise. Still, since the study is archival, we cannot precisely determine the extent to which players experienced pressure and/or explicitly monitored their skills concurrently. Due to this limitation, the theoretical nature of the study's conclusions must be interpreted with a certain level of caution.

Along these lines, not all at bats and innings pitched *within* the season and postseason are quite alike. Perhaps an at bat in the early innings of a regular season contest at the start of the year inspires a different level of anxiety than one in a tie game late in the year; likewise, an inning pitched in a Division Series game might feel different than one in the World Series. Specific regular season appearances might even inspire greater pressure than some of those from the postseason. Thus, a recommended follow-up study might break down our data further; a pressure ranking could be established for each at bat or inning pitched on the books, for example, and this variable could then be compared to performance. For this, that at bat early in the year would rank near the bottom on pressure, for instance, while an inning pitched in the World Series might rank a dozen or so intervals higher.

Another logical extension of this research is the application to other sports. Public, historical records of sporting events and leagues are kept and updated every day, and pressure could be similarly operationalized within each for analysis. Hypotheses might vary depending on the skills involved in the sport; for example, for basketball and American football popular wisdom often recommends that "defense wins championships" (see Moskowitz & Wertheim, 2012); this claim could conceivably be tested.

This returns us to our central individual-level hypothesis for baseball, that hitters are more prone to variance in explicitly monitoring their skills under pressure, since hitting is presumably a most complex skill (as in Kinrade et al., 2010; Masters et al., 1993). This prediction was met with modest support; batting average was indeed our least robust statistic when crossing from the regular season to the postseason, but all other hitting stats correlated well. Thus, findings suggest that the act of hitting a baseball is a bit more sturdy under pressure than we thought, but that psychological factors may impact hitting the ball for contact (e.g., as measured by batting average) more so than for power (e.g., as measured by OPS). Based on this, for training under pressure we might recommend a focus on contact hitting, paired with a maintenance of trust in one's existing skills and experience when hitting for power. The latter could be then be extended to skills of base-stealing and pitching, as we had originally projected; if one is successful under less pressure, then he should be confident that those skills may also translate to high-anxiety conditions.

For teams, data from recent years have supported the position that good regular season pitching should lead to postseason success, over and above regular season hitting. This has not always been the case across the history of baseball, however, leaving us to speculate on changes along the way in how the game has been played. Meanwhile, reuniting this with our player-level hypotheses leads us to conclude that pitching does seem relatively robust when translated from lesser to greater pressure, and in recent history helps team postseason performance accordingly. Good hitting, on the other hand, while perhaps more robust than originally projected, does not as much lead to successful performance under pressure, at least not recently.

Thus, what might we advise a general manager hoping to build the next successful, clutch baseball squad? The data suggest to build around pitching first, but to also not neglect hitting entirely, since season power numbers seem also to correlate well to the postseason at the individual level. Future research will help build upon these results, in particular baseball-specific empirical studies such as that of Gray (2004).

Acknowledgments

For help with the arduous tasks of data collection and management, thanks to researchers Johnny Felt, Marc Grover, Stefanee Maurice, Travis Miller, Mimi Okojie, Michael Overstreet, Robby Rhodes, Tiffany Widjaja, Jenna Yowell and Rocky Zamora.

References

Baseball-Reference.com. (2012). *Playoff and world series stats and results*Available from http://www.baseball-reference.com/postseason/.

Baumeister, R. F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 16*, 610–620. http://dx.doi.org/10.1037/0022-3514.46.3.610.

Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: choking under pressure in sports and mental tests. *European Journal of Social Psychology, 16*, 361–383. http://dx.doi.org/10.1002/ejsp.2420160405.

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: what governs choking under pressure? *Journal of Experimental Psychology: General, 130*, 701–725 http://dx.doi.org/10.1037/0096-3445.130.4.701.

Cheng, W.-N. K., Hardy, L., & Markland, D. (2009). Toward a three-dimensional conceptualization of performance anxiety: rationale and initial measurement development. *Psychology of Sport & Exercise, 10*, 271–278. http://dx.doi.org/10.1016/j.psychsport.2008.08.001.

DeCaro, M. S., Thomas, R. D., Albert, N. B., & Beilock, S. L. (2011). Choking under pressure: multiple routes to skill failure. *Journal of Experimental Psychology: General, 140*, 390–406. http://dx.doi.org/10.1037/a0023466.

Gray, R. (2004). Attending to the execution of a complex sensorimotor skill: expertise differences, choking, and slumps. *Journal of Experimental Psychology: Applied, 10*, 42–54. http://dx.doi.org/10.1037/1076-898X.10.1.42.

Jackson, R. C., Ashford, K. J., & Norsworthy, G. (2006). Attentional focus, dispositional reinvestment, and skilled motor performance under pressure. *Journal of Sport & Exercise Psychology, 28*, 49–68.

Kinrade, N. P., Jackson, R. C., Ashford, K. J., & Bishop, D. T. (2010). Development and validation of the decision-specific reinvestment scale. *Journal of Sports Sciences, 28*, 1127–1135. http://dx.doi.org/10.1080/02640414.2010.499439.

Lowry, R. (2012). *Two correlation coefficients*. Retrieved from. http://vassarstats.net/rdiff.html.

Masters, R. S. W. (1992). Knowledge, knerves and know-how. *British Journal of Psychology, 83*, 343–358.

Masters, R. S. W., Polman, R. C. J., & Hammond, N. V. (1993). 'Reinvestment': a dimension of personality implicated in skill breakdown under pressure. *Personality and Individual Differences, 14*, 655–666. http://dx.doi.org/10.1016/0191-8869(93)90113-H.

Merron, J., & Schoenfield, D. (2005). *ESPN.com: page 2: Playoff theories: Do they hold up?*. Retrieved from. http://sports.espn.go.com/espn/page2/story?page=baseball/one/051003.

Mesagno, C., Harvey, J. T., & Janelle, C. M. (2011). Self-presentation origins of choking: evidence from separate pressure manipulations. *Journal of Sport & Exercise Psychology, 33*, 441–459.

Moskowitz, T., & Wertheim, L. J. (2012). *Scorecasting: The hidden influences behind how sports are played and games are won*. New York, NY: Three Rivers Press.

Otten, M. P. (2009). Choking vs. clutch performance: a study of sport performance under pressure. *Journal of Sport & Exercise Psychology, 31*, 583–601.

Peterson, D. (2011). Hitting a baseball – "The hardest thing to do in sports". *Axon Sports*. Retrieved from. http://www.axonpotential.com/hitting-a-baseball-the-hardest-thing-to-do-in-sports/.

Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76*(2), 92–104. http://dx.doi.org/10.1037/h0031332.