Basic and Applied Social Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hbas20

Evidence of Choking Under Pressure on the PGA Tour

Brett M. Wells [a] & John J. Skowronski [a]
[a] Northern Illinois University

Version of record first published: 06 Apr 2012.

To cite this article: Brett M. Wells & John J. Skowronski (2012): Evidence of Choking Under Pressure on the PGA Tour, Basic and Applied Social Psychology, 34:2, 175-182

To link to this article: http://dx.doi.org/10.1080/01973533.2012.655629

BASIC AND APPLIED SOCIAL PSYCHOLOGY, 34:175–182, 2012

ISSN: 0197-3533 print/1532-4834 online
DOI: 10.1080/01973533.2012.655629

Psychology Press
Taylor & Francis Group

Evidence of Choking Under Pressure on the PGA Tour

Brett M. Wells and John J. Skowronski
Northern Illinois University

The authors examined whether there was evidence of choking under pressure (CUP) among professional golfers on the 2009 PGA Tour. Following the suggestion of Beilock and Gray (2007), choking was determined via within-golfer comparisons. Analyses yielded strong evidence of CUP in that evidence of performance decline was greatest when pressure was greatest. That is, across a span of 28 years, 4th-round tournament scores were significantly worse than 3rd-round tournament scores. Moreover, the magnitude of the choking effect was related to a player's position on the leaderboard: The closer a player was to a tournament lead, the larger his choking score. Finally, the nature of the analyses conducted makes it unlikely that the obtained effects can be solely attributed to statistical phenomena such as regression to the mean.

> "When the body fails or is weary, and that shot into a howling gale seems impossible, it is the mind that enables a player to pull it off." — Gary Player (2010, p. ix; Winner of 165 golf tournaments worldwide, including nine majors on the regular PGA Tour and nine majors on the Senior PGA Tour)

THE GOLFING ROUNDS OF PHIL MICKELSON: CHOKER OR CLUTCH PERFORMER?[1]

At the 2010 Masters Tournament, Phil Mickelson rallied and shot a 32 (4 under par[2]) on the last nine holes to win the tournament. This "back-nine" performance has become noted for one especially outstanding shot. This occurred on the 13th hole, a 480-yard par 5 named "Azalea" on which a water hazard, "Rae's Creek," guards the green. The late Robert "Bobby" Jones (1959; golfing icon, designer of Augusta, and originator of the Master's Tournament) eloquently summarized the decision that golfers have to make on this hole:

> Whatever position may be reached with the tee shot, the second shot as well entails a momentous decision whether or not to try for the green. Several tournaments have been won or lost here, even though the decision may not have been obvious at the time. (p. 230)

Indeed, Phil probably faced a more daunting decision than the one envisioned by Mr. Jones. Not only did Phil's drive leave his ball a long distance from the hole and sitting among trees in pine straw, but Phil also had to thread his second shot between two imposing trees. In this instance, with a major championship on the line, Phil hit a superb shot. He precisely placed his golf ball within 10 feet of the hole. Reflecting on his improbable shot some 6 months later, Phil referred to it as the greatest of his career, "not because it was the hardest shot or the best shot but because it came at such a clutch time and ultimately led to victory in my favorite event" (Shipnuck, 2010, p. 110).

Although it is easy to find instances in which Phil's performance was *clutch*, it is seemingly just as easy to

[1]We don't want to appear to either promote or denigrate Phil Mickelson. The selection of Phil Mickelson as the lead personality in this article was based on the fact that his golf performances exhibit both stunning successes and stunning failures. As such, among those on the current golf tour, he was the perfect individual to use as an example of someone who produces both seeming clutch performances and seeming occurrences of CUP.

[2]Par is a term in the game of golf to denote the number of strokes that a golfer should require to finish a hole (i.e., 3, 4, or 5). If a golfer requires fewer strokes to complete a hole, he or she is under par; if a golfer requires more strokes to complete a hole, he or she is over par. Summing the par values for each of the 18 holes gives the par for the course. Augusta National Golf Club, like most championship golf courses, is a par 72 comprised of a front nine and back nine holes (i.e., first nine and second nine holes, respectively). Both the front nine and back nine holes are par 36 (i.e., half of 72). Hence, Phil Mickelson required four fewer strokes than par to complete the back nine holes.

Correspondence should be sent to Brett M. Wells, Department of Psychology, Northern Illinois University, DeKalb, IL 60115. E-mail: bwells1@niu.edu

find instances when Phil seemed to *choke under pressure* (CUP; Beilock & Carr, 2001; Lewis & Linder, 1997; Masters, 1992). In a notable example, Phil started the 72nd hole of the 2006 U.S. Open with a one-shot lead over Geoff Ogilvy. Phil double-bogeyed the hole, giving Ogilvy an unlikely one-shot victory. Ogilvy was utterly stunned at Phil's collapse: "I think I was the beneficiary of a little bit of charity." Sadly, not even Phil himself could muster a suitable response regarding his self-destruction: "I am still in shock that I did that. I am such an idiot. I just couldn't hit a fairway all day. I just can't believe I couldn't par the last hole. It really stings" (Leonard, 2006).

CHOKING, CLUTCH, OR ILLUSION?

Sportswriters and fans make statements about choking and clutch performances all the time, and players certainly express belief in such phenomena (see Gucciardi, Longbottom, Jackson, & Dimmock, 2010). However, in the scientific world the labels *choke* and *clutch* are a bit controversial. These labels imply that something out of the ordinary occurred that caused the person's performance to differ from that person's normal performance level. An alternative view is that such events simply reflect the normal ups and downs of a person's performance, but it is the magnitude of the situation that causes perceivers to make inappropriate inferences about the person as having choked or as having been clutch (an idea derived from Gilovich, Vallone, & Tversky, 1985).

Certainly, some analyses of real-world performance data are consistent with the notion of out-of-ordinary performances. For example, Schell (2005) conducted an extensive statistical analysis of baseball data and concludes that some hitters (e.g., Yogi Berra) exhibited a performance profile that warranted the label "clutch hitter." Similarly, data from Baumeister and Steinhilber (1984); Dohmen (2008); Wright, Jackson, Christle, McGuire, and Wright (1991); and Wright, Voyer, Wright, and Roney (1995) suggested that athletes who play in front of a supportive audience perform less successfully than visiting players when they have the chance to capture a championship.

However, other data are inconsistent with the notion that athletes are chokers or clutch performers (see Gayton, Matthews, & Nickless, 1987; Schlenker, Phillips, Boniecki, & Schlenker, 1995a, 1995b). For example, Clark (2002a) used data from real Professional Golf Association (PGA) golf tournaments to examine the hypothesis that the greatest levels of pressure should be on golfers who were near the top of the leaderboard and that such pressure should be reflected in increases (poorer performance) in golfers' final round scores. This, of course, is because pressure should be greatest for those in a position to win the tournament: A win on the PGA Tour entails a hefty paycheck (usually in the amount of $1 million or more), exemptions to play in major championships (tournaments set aside for the most elite group of professional and amateur golfers), and a degree of job security (if a player wins a PGA Tour event he cannot lose his tour card for the following two seasons). Failure to win yields a substantially smaller paycheck and typically provides none of the other direct benefits.

Clark's (2002a) results suggested that golfers who were leading a tournament or who were one shot from the lead did not have significantly higher final round scores than those golfers who were two to five shots from the lead. Moreover, he noted that golfers who were leading going into the final round of play won the majority of the time. Clark (2002b) later replicated these findings in a study examining data from the PGA Tour's qualifying tournaments in which the competitors are golfers on the verge of becoming PGA Tour professionals. Using such results, Clark concluded that there was insufficient evidence to support the assertion that professional golfers exhibited CUP. This qualifying tournament null outcome was especially impressive given that there is a great deal riding on a player's tournament result (e.g., money, endorsements, job security, fame, etc.), so the qualifying tournaments are regarded as especially stressful and pressure packed.

Such null results are surprising to those who conduct laboratory research into the phenomenon of choking under pressure. The evidence from these laboratory studies reliably shows that experts in a variety of domains often exhibit poor performance in pressure-packed situations (Baumeister, 1984; Beilock & Carr, 2001; Beilock, Carr, MacMahon, & Starkes, 2002; Gray, 2004; for a review, see Beilock, 2010). Indeed, one might expect that the ubiquitous CUP effects found in the laboratory would be amplified in real-life situations in which the stakes, and thus the pressure to perform, are especially high (Beilock & Gray, 2007). Thus, the absence of evidence in Clark's analyses indicative of a choking effect is a bit puzzling.

Several hypotheses might be offered for Clark's null effect findings. It could be that the effect is simply a small one, reliably detectable in the laboratory (where extraneous variance is minimized) but less detectable in the high-variance real world. An alternative possibility is that choking might be exhibited on the kinds of tasks employed in laboratory contexts but that real-world tasks are less subject to the effect (for a similar idea in the realm of media effects on aggressive behavior, see Savage & Yancey, 2008). However, Beilock and Gray (2007) suggested a third idea: that Clark's failures to obtain evidence of CUP might lie in his use of

between-golfer comparisons and his reliance on final-round scores. That is, one of the major limitations of Clark's conceptualization of CUP is that it does not take into account variation within golfers—it exclusively measures variation between golfers. Beilock and Gray discussed this notion when they suggest an alternative approach to these data:

> One might compare a golfer's initial-round score (or an average over a few initial or nonpressure rounds) with his or her final round "pressure" score. Golfers with a higher score in the final round than in their initial rounds could be thought of as choking. (p. 428)

There is an obvious statistical advantage to conducting the analyses in a within-subject manner. By using golfers as their own controls, one can achieve additional control over extraneous variance and, at least in theory, conduct statistical tests that are especially sensitive to the presence of the effect under investigation. Moreover, according to Beilock and Gray (2007), because CUP is often defined as performing worse than expected as compared to one's *own* typical performance level (not the performances exhibited by others), such analyses also make more theoretical sense than Clark's.

To highlight the issue of between- versus within-golfer variation, consider the following example: Imagine that two golfers enter the final round of play and Golfer A is leading the tournament over Golfer B by a single shot. During the final round of the tournament, the two golfers shoot identical scores (e.g., 72), and Golfer A wins the tournament. According to Clark's approach (examination of between-golfer variation), Golfer A would not be CUP—after all, he did win the tournament and carded the same score as Golfer B. However, what if you were presented with information about each golfer's typical performance (examination of within-golfer variation)? It is indeed possible that Golfer A could be performing worse than his average (e.g., 70.5) and Golfer B could be performing better than his average (e.g., 73.5). According to Beilock and Gray (2007), despite the fact that he won the tournament, in such a situation, Golfer A may be exhibiting CUP.

SUMMARY AND OVERVIEW

In summary, there has been controversy as to whether CUP is exhibited by professional athletes, despite the fact that such effects are often demonstrated and replicated in laboratory settings. Beilock and Gray (2007) suggested that this disconnect between analyses of the performances of highly skilled athletes in field studies and in laboratory experiments may, at least in part, rest in a misconceptualization of CUP. That is, instead of defining CUP as a between-athlete phenomenon, they suggest defining it as a within-athlete phenomenon. To this end, we used Beilock and Gray's within-golfer conceptualization of CUP to examine whether evidence for CUP emerged from archival performance data obtained from golfers on the PGA Tour.

METHOD

Subjects

The subjects in the study were professional members of the 2009 PGA Tour. The regular season of the 2009 PGA Tour consisted of 38 tournaments. Performance data from 34 tournaments were used for the study. Most PGA Tour tournaments contain four rounds of golf (with many cutting the field after two rounds) and are medal (stroke) play.[3] Because they were deemed to significantly deviate from this standard, four tournaments were excluded from analyses: One tournament (The 50th Bob Hope Classic Hosted by Arnold Palmer) required five rounds of play, one tournament (The PGA Tour Qualifying Tournament) required six rounds of play and does not include the vast majority of PGA Tour professionals, one tournament (AT&T Pebble Beach National Pro-Am) involved only three rounds due to inclement weather, and another tournament (World Golf Championships: Accenture Match Play Championship) involved match play.

Although Clark (2002a, 2002b) limited his analyses to golfers who were within five shots of the leader(s) entering the final round of play, we chose to include all golfers who made the cut following the second round of play. One reason for this decision is statistical: Limiting analyses to only those golfers within five shots of the lead dramatically restricts the range of final round scores of an already homogeneous sample of golfers (Clark, Woodward, & Wood, 2008). It is well known that such restriction of range can obscure evidence of an effect that is actually present in the data. A second notion underlying this decision, though, was theoretical. We hypothesized that pressure is related to positioning on the leaderboard in a continuous fashion, with those at the top of the leaderboard feeling the most amount of pressure and those at the bottom feeling the least amount of pressure.

[3]After the second round of play, the field of golfers is reduced (i.e., "cut") to those with the lowest 70 tournament scores (and ties). Medal (stroke play) is the scoring system for most of the tournaments on the PGA Tour. It involves summing a player's strokes on each hole across the four rounds of play. The winner is the golfer with the lowest number of strokes.

Procedure

Players' scores for each round of the 34 tournaments were harvested from http://www.pgatour.com. In addition, the number of shots behind the leader(s) was calculated for each player at the end of the first, second, and third rounds. For example, at the end of the third round of the 2009 Northern Trust Open, Phil Mickelson was alone in first place with a total score of 197. Ernie Els, on the other hand, was tied for 51st place with a total score of 211. Ernie, therefore, was 14 shots behind the leader, Phil, when starting the final round of play.

CUP scores. Following the suggestions of Beilock and Gray (2007), the within-golfer choking score was operationalized by comparing a golfer's final- (fourth) round score with his third-round score. Thus, choking is defined by the following equation: Choking = (Round Round 4 Score) – (Round 3 Score). Positive scores may reflect choking.[4]

RESULTS

Initial Evidence of CUP

Player choking scores were significantly greater than zero ($M = .43$, $SD = 3.99$), $t(2038) = 4.92$, $p < .001$, Cohen's $d = .22$. In addition, there was a strong correlation between a player's choking score and the number of shots that he was off the lead when starting the fourth round, $r(2039) = -.440$, $p < .001$. That is to say, the fewer shots a player was from the lead when starting the fourth round, the higher the choking score. Both effects are consistent with the CUP hypothesis.

The CUP Even Appears in Aggregate Difference Scores

The analyses reported in the preceding paragraph are potentially contaminated by the fact that the data points are not independent: A single golfer could contribute multiple data points to the analysis. This violates the independence-of-observations assumption of the t test that was used to examine the data. One way to circumvent this problem is via averaging. Accordingly, all round scores were averaged separately by player so that each player contributed only one difference score (season average for Round 4 – season average for Round 3) to the analysis. Again, the presence of a CUP is indicated if the difference scores are greater than 0. This effect was present in the analysis ($M = .48$, $SD = 1.31$), $t(188) = 5.00$, $p < .001$, Cohen's $d = .73$. Moreover, we again observed a strong correlation between a player's average choking score and his average number of shots from the lead when starting the fourth round, $r(189) = -.409$, $p < .001$.

Is the CUP Effect Limited to Data From the 2009 PGA Tour?

One might question the generality of the observed CUP effects, asking whether the effect was simply an anomaly or whether it consistently appeared across years. To examine this idea, third-round averages and fourth-round averages, averaged across players and across tournaments from 1983 until 2010, were harvested from the PGA Tour website (http://www.pgatour.com). Important to note is that this data file reflected averages from every PGA Tour player and tournament in a given season. The fact that results were averaged across these variables means that data included in this analysis were not identical to those data previously described. For example, analyses included data from PGA Tour events that fall outside of the regular PGA Tour season (e.g., the fall series and the newly created PGA Tour Playoffs). Moreover, these data are a bit sloppy in that for a very few tournaments in each year, the tournament's third and fourth rounds were not the final two rounds of play.

Nonetheless, analyses of these data suggest that in the past 28 years on the PGA Tour there is consistent evidence suggestive of CUP (see Table 1). PGA Tour players averaged .32 strokes higher during their fourth rounds than their third rounds, $t(27) = 8.74$, $p < .001$, Cohen's $d = 3.36$. Of interest, these effects also replicated for the year in which Clark (2002a) conducted his analysis (the 1999 PGA Tour). In that year, players' CUP scores (Fourth-Round Average – Third-Round Average) were significantly greater than zero ($M = .53$, $SD = 1.15$), $t(191) = 6.35$, $p < .001$, Cohen's $d = .918$.

Can the CUP Effects Be Explained by Regression to the Mean?

One concern in these analyses is whether these CUP effects can be simply explained by the statistical phenomenon of regression to the mean. This phenomenon is based on the notion that a player has a stable, long-term scoring average but that the player's scores exhibit random variation around that average. Because scores near the average score are most likely, exceptional scores should tend to be followed by average scores. Random chance thus suggests that players who post

[4]We conducted other analyses defining choking scores in alternative ways. For example, one might look for evidence of choking by calculating deviations from a player's season-long scoring average. The magnitude of the choking effects that were observed did not substantially vary by the way in which the choking score was determined. Because of this redundancy, we chose to not report results from these alternative analyses; interested readers may contact the first author for results.

TABLE 1
Choking on the PGA Tour From 1983 Until 2010

Year on the PGA Tour	*Round 4 Average*	*Round 3 Average*	*Choking*
2010	71.10	70.44	0.66
2009	70.88	70.73	0.15
2008	71.20	70.91	0.29
2007	71.42	71.18	0.24
2006	71.45	71.01	0.44
2005	71.45	70.93	0.52
2004	71.60	71.02	0.58
2003	70.98	70.80	0.18
2002	71.11	71.02	0.09
2001	70.95	70.73	0.22
2000	71.02	70.95	0.07
1999	71.99	71.47	0.52
1998	71.35	71.17	0.18
1997	71.67	71.31	0.36
1996	71.35	70.95	0.40
1995	71.38	71.18	0.20
1994	71.51	71.00	0.51
1993	71.48	71.34	0.14
1992	71.18	71.00	0.18
1991	71.74	71.03	0.71
1990	71.46	71.57	−0.11
1989	71.47	71.09	0.38
1988	71.26	70.97	0.29
1987	71.66	71.30	0.36
1986	72.06	71.59	0.47
1985	72.34	71.76	0.58
1984	72.11	71.94	0.17
1983	72.14	71.83	0.31
M	71.48	71.15	0.32
SD	0.38	0.35	0.20

Note. Choking is equal to fourth-round scoring average minus third-round scoring average. These values reflect only PGA Tour professionals and not those who might have qualified for a tournament. Moreover, these values reflect all PGA Tour tournaments.

an above-average score in the third round of a tournament might thus be expected to score more poorly in the tournament's fourth round.

To address the regression to the mean idea, additional regression analyses were conducted on the 2009 PGA Tour player performance data. Key to these analyses is the idea that the regression to the mean argument is one of *equal opportunity,* in that it suggests that the same amount of regression should occur *regardless of the round*. Hence, if a golfer shoots a 62 in the second round, he should experience the same amount of regression to the mean in the third round as if he shot a 62 in the third round and experiences regression to the mean in the fourth round. Thus, if one could find *greater decrements of performance in the fourth round* as compared to earlier rounds, then one would have evidence that at least part of the fourth-round decrease in performance is not due to regression to the mean—and could be caused by choking.

Using this idea, we calculated score differences for both the third round relative to the second round (Round 3 – Round 2) and for the fourth round relative to the third round (Round 4 – Round 3). If the fourth-round performance decrease previously described was simply a regression to the mean effect, a performance decrease of a similar magnitude should be observed when comparing performance in Round 3 to performance in Round 2. If the effect that we reported in the earlier analyses is not a simple regression to the mean effect, but instead reflects choking, then the performance decrease should be greater for Round 4 relative to Round 3 than for Round 3 relative to Round 2.[5]

The following variables were included in regression analyses examining these difference scores: PGA Tour player (so these analyses again focus on PGA Tour pros only), tournament, standing (the number of shots behind the leader), and round (third vs. fourth). In the first step of these analyses, these variables were regressed on the scoring differences (Third-Round Score – Second-Round Score and Fourth-Round Score – Third-Round Score) exhibited by those players who made the cut in each tournament. A second step in the regression analysis included all these terms, as well as the Standing × Round interaction.

Note that the analyses that we performed were pooled hierarchical within-subjects regression analyses in which, much as in mixed-model analyses of variance, the between-golfer variance is partialled from the within-golfer variance prior to examining the within-golfer effects. Thus, these analyses do not violate independence-of-observation assumptions: The inclusion of multiple difference scores from the same golfer in the analysis is entirely legitimate.[6]

[5]The difference scores that we calculated (Round 4 – Round 3; Round 3 – Round 2) are a conceptually correct test of the regression to the mean idea. That is, conceptually, a test of the regression idea would first involve a comparison of a player's round score to the player's long-term mean, and then would compare these difference scores across rounds. However, given that the season means are nearly constant in all terms, one can eliminate them from the difference score calculations. Thus, the difference scores that we calculated and compared do indeed provide an assessment of the regression to the mean idea.

[6]This analysis also allows us to calculate the CUP scores on a golfer-by-golfer basis. We have chosen not to report any of the names of the identified "chokers." Of interest, however, we will note that Tiger Woods averaged .44 strokes *lower* in his fourth rounds of 2009 when compared to his total season scoring average. In addition, in 2009 Tiger's finish position improved or remained unchanged in the final round 93.33% of the time (Ranked 1st on tour). In comparison, during Tiger's tumultuous 2010 season, he averaged .33 strokes higher in his final rounds as compared to his scoring average. The year 2010 was also Tiger's first year since turning professional without recording a victory on the PGA Tour. At first glance, it appears that the performances of one of the best players on tour, and perhaps one of the greatest golfers of all time, was susceptible to the effects of stress and pressure.

We expected that evidence supporting the notion of CUP should come from two sources. One source is the round: Pressure should be greater in Round 4 of a tournament than in Round 3. The second is the position occupied by a player on the leaderboard at the start of a round: The higher the standing, the higher the pressure. We are agnostic as to whether these effects are additive or interactive: Both patterns can support the theory. That is, the choking idea would be supported by data showing that choking was enhanced as a result of the additive consequences of two main effects (one for round and one for leaderboard position). It would also be supported by an interaction indicating that the amount of choking among the leaders in Round 4 of a tournament was greater than the amount expected from the additive combination of the round and leaderboard position main effects.

The results of the analyses examining the between-subjects effects were straightforward. The results suggested that the difference scores varied across players, $F(188,\ 3850) = 1.27,\ p = .008$. Likewise, the difference scores were significantly greater for some tournaments than for others, $F(33,\ 3850) = 8.15,\ p < .001$.

More critical were results involving the standing variable. The standing effect could reflect the CUP hypothesis: The higher a player's placement in the standings at the start of a round, the greater the felt pressure and the greater the performance reduction relative to the previous round. In accord with this idea, the standing effect was significant, $F(1,\ 3850) = 993.99,\ p < .001,\ \beta = -.540$.

One issue with the standing effect is that it could reflect regression to the mean: A higher standing is likely related to a superior performance in the prior round, and the regression idea suggests that a performance that is worse than the previous superior performance is likely in the subsequent round. This interpretive problem would not apply to a round main effect. Only the choking hypothesis, and not the regression to the mean hypothesis, suggests that the pressure should be greater in Round 4 than in Round 3, so that the performance decrease between a round and the prior round should also be greater in Round 4 than in Round 3. Indeed, the round effect was significant, $F(1,\ 3850) = 23.23,\ p < .001,\ \beta = -.540$, and the means are consistent with the CUP hypothesis (Least Squares Means: Round 3 – Round 2 = .62, Round 4 – Round 3 = 1.16). This difference seems small, but in the context of PGA Tour performance standards, it is fairly substantial. For example, in 2007 a reduction in performance of .54 strokes per round was the difference between the 19th best tour scoring average (Adam Scott, 70.01) and the 57th best average (Fred Funk, 70.55).

The Standing × Round interaction also could provide evidence of a choking effect. Such an interaction could show that choking was especially great in Round 4 among those highest in the standings at the end of Round 3. However, this interaction did not emerge, $F(1,\ 3849) = 0.74,\ p = .391$. Nonetheless, these data are still consistent with the CUP hypothesis. The absence of an interaction simply indicates that the standing and round effects are additive and are consistent with the notion that the greatest pressure occurs when one is both in the fourth round and at the top of the leaderboard at the start of the round.

Readers may wonder whether the effects that we report were dependent on the rounds that were used to calculate the difference scores. This is not the case. We conducted a number of analyses of the performance data, and results were consistent across analyses: Relative to other rounds, performance decreased in Round 4 compared to all other rounds (details available on request).

DISCUSSION

Beilock and Gray (2007) suggested that it would be fruitful to use within-subject analyses to search for evidence that professional players CUP. Such analyses make sense both from the perspective of theory (choking refers to a decrease in performance relative to a player's normal level) and statistically (within-subject analyses are especially sensitive to the presence of effects). It appears that Beilock and Gray were correct: In the present study, this evidence suggestive of CUP emerged from analyses of performance data for the 2009 regular season of the PGA Tour, appeared across different types of analyses, was replicated in analyses of data from performances over the past 28 years on tour, was observed in the same data set examined by Clark (2002a) that reported no such effects in a between-subjects analysis, and cannot be fully explained by regression to the mean effects. In fact, we note that it is also likely that the presence of practice effects underestimates the obtained CUP effects. That is, over the course of a tournament, professional golfers may become better accustomed to the course conditions (e.g., green speed, green hardness, approach angles, etc.) and might consequently be expected to play better as they play additional rounds at the tournament course. Plainly, in the course of a tournament, they do not—especially in the final round when there is a chance of winning the tournament.

Limitations

The analyses conducted in this study document performance decrements. Although such decrements are consistent with the idea of CUP, other mechanisms might apply. For example, performance decrements in the fourth round might be due to player fatigue (physically,

mentally, or both) instead of CUP. However, although fatigue might partly contribute to performance decrements in the final round, experimental data from Jackson, Ashford, and Norsworthy (2006) demonstrate CUP effects when controlling for the impact of fatigue.

Nonetheless, it is extremely difficult to disentangle such notions via reliance on archival data. Hence, future research that wishes to more closely examine the reasons underlying the performance decrements that we document may have to conduct direct research. Such research may attempt to probe the internal psychological states of golfers during a tournament, or might retrospectively ask golfers to report on their mental and physical states as a tournament progresses. Such data might be best combined with observational data that assesses the variability in players' pre-shot routines. Indeed, Beilock, Bertenthal, McCory, and Carr (2004) provided initial evidence that experienced golfers under pressure putted best when they spent less time over a putt. This finding is quite the contrary to the popular notion of "trying harder" or "concentrating longer" over critical golf shots. In addition, retrospective recollections might be able to examine the mechanisms thought to prevent CUP (i.e., changes in positive self-talk, relaxation, imagery, pre-shot routines, focus, etc.).

Others who examine our results may criticize them because they do not control for factors such as inclement weather, including rain, wind, and poor visibility. Obviously, these can all lead to inflated scores. However, such factors are randomly distributed across the rounds of play and therefore should not substantially affect our choking estimates—especially those examining results for the past 28 years on the PGA Tour.

Similarly, one may question whether the course conditions themselves are randomized across the rounds of play (e.g., pin placement, green speed, rough length, etc.) or whether the superintendents of golf courses alter these conditions to make the final round of play especially difficult in hopes of creating an extremely entertaining final round of play. However, anecdotal evidence suggests that much care is taken to assure that hole difficulty is relatively equivalent across the rounds of play. For example, Tom Meeks, a former United States Golfing Association Senior Director of Rules and Competition, dismissed the popular notion of *Sunday Flags*—intentionally placing holes in the toughest possible location during the final round of play (Feinstein, 2003). Instead, Meeks noted that to aid his process of selecting hole locations, he would first pick four hole locations on each green and rank them in difficulty. Meeks would then distribute these hole locations across the four rounds of play so that the sum of hole difficultly rankings would be as close to 45 as possible (i.e., 2.5×18). Indeed, as Meeks stressed, "You don't want to start out with the eighteen easiest hole locations on Thursday and have the eighteen hardest ones on Sunday. ... You want balance" (Feinstein, 2003, p. 217). Consequently, we believe that performance decrements in the fourth round cannot easily be attributed to sudden changes in course difficulty.

Nonetheless, such considerations emphasize that our archival analyses only document performance decrements and that those performance decrements can be caused by many mechanisms. It is in this regard that the laboratory research conducted on this topic (e.g., Beilock, 2010; Beilock & Carr, 2001; Beilock et al., 2004) is exceptionally useful, for that research not only documents the reality of the choking under pressure phenomenon but also documents some of the mental mechanisms that contribute to the effect.

Future Directions

Our analyses focused exclusively on players' scores and not other metrics of performance. As the popular, grill-room adage stresses, "The scorecard doesn't ask how, but rather how many." There are many aspects to performing well in golf, including (but not limited to) driving, iron play, short game, and putting. Future research might examine whether pressure affects all aspects of the game equally or whether certain aspects of the game are more susceptible to pressure than others (e.g., the yips in putting; Smith et al., 2003; Smith et al., 2000).

For example, researchers may utilize data from Shot Tracker™ (www.pgatour.com/shottracker), a system that records hole-by-hole data for each player. One metric of importance may include average distance to the pin. Such a measurement is perhaps the purest estimate of accuracy. Based on the data presented in this article, one might predict that those who are especially prone to CUP (i.e., those near the top of the leaderboard) would be less accurate with their approach shots than their previous rounds and/or average rounds might predict. Other metrics of interest may include the total feet of putts made, three-putt avoidance, driving distance, and driving accuracy. Moreover, utilizing such data may help to determine when CUP is likely to occur: at the beginning of the final round, the end of the final round, or throughout the duration of the final round.

It is also important to note that researchers should extend Beilock and Gray's (2007) within-athlete comparisons to other sports (e.g., batting for baseball players [Schell, 2005] or shooting percentage for basketball players). Such analyses are obviously important if CUP is to be shown to be a general phenomenon that emerges across sporting domains. Last, meta-analyses examining CUP in real-word game situations may examine whether the conceptualization of CUP (between- vs. within-athlete) moderates CUP effect sizes. Such an examination would help clarify the

conditions under which CUP is most clearly demonstrated. We hope that future research will shed light on these issues so that theorists can better understand whether CUP is, indeed, a real and demonstrable phenomenon or whether it is simply a function of the overexuberant inferential processes of fans of professional and amateur sports.

ACKNOWLEDGMENTS

An earlier draft of this article was presented as a poster at the Association for Psychological Science's 22nd Annual Convention in Boston, Massachusetts, May 28, 2010.

REFERENCES

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, *46*, 610–620. doi: 10.1037/0022–3514.46.3.610

Baumeister, R. F., & Steinhilber, A. (1984). Paradoxical effects of supportive audiences on performance under pressure: The home field disadvantage in sports championships. *Journal of Personality and Social Psychology*, *47*, 85–93. doi: 10.1037/0022–3514.47.1.85

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

Beilock, S. L., Bertenthal, B. I., McCory, A. M., & Carr, T. H. (2004). Haste does not always make waste: Expertise, direction of attention, and speed versus accuracy in performing sensorimotor skills. *Psychonomic Bulletin and Review*, *11*, 373–379. doi: 10.1037/a0012859

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. doi: 10.1037/0096–3445.130.4.701

Beilock, S. L., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied*, *8*, 6–16. doi: 10.1037/1076–898X.8.1.6

Beilock, S. L., & Gray, R. (2007). Why do athletes "choke" under pressure? In G. Tenebaum & R. C. Eklund (Eds.), *Handbook of sport psychology* 3rd ed., (pp. 425–444). Hoboken, NJ: Wiley.

Clark, R. D., III. (2002a). Do professional golfers "choke"? *Perceptual and Motor Skills*, *94*, 1124–1130. doi: 10.2466/PMS.94.3.1124

Clark, R. D., III. (2002b). Evaluating the phenomenon of choking in professional golfers. *Perceptual and Motor Skills*, *95*, 1287–1294. doi: 10.2466/PMS.95.7.1287

Clark, R. D., III, Woodward, K. L., & Wood, J. M. (2008). On the unreliability of golf scores for professional golfers: A case for restriction of range. *Perceptual and Motor Skills*, *107*, 683–690. doi: 10.2466/PMS.107.7.683

Dohmen, T. J. (2008). Do professionals choke under pressure? *Journal of Economic Behavior and Organization*, *65*, 636–653. doi: 10.1016/j.jebo.2005.12.004

Feinstein, J. (2003). *Open: Inside the ropes of Bethpage Black*. Boston, MA: Little, Brown.

Gayton, W. F., Matthews, G. R., & Nickless, C. J. (1987). The home field disadvantage in sports championships: Does it exist in hockey? *Journal of Sports Psychology*, *9*, 183–185.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, *17*, 295–314.

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. doi: 10.1037/1076–898X.10.1.42

Gucciardi, D. F., Longbottom, J., Jackson, B., & Dimmock, J. A. (2010). Experienced golfers' perspectives on choking under pressure. *Journal of Sport and Exercise Psychology*, *32*, 61–83.

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

Jones, R. T., Jr. (1959). *Golf is my game*, Garden City, NY: Doubleday.

Leonard, T. (2006, June 19). I am such an idiot: Mickelson badly misplays final hole to present title to Ogilvy. *U-T San Diego*. Retrieved from http://www.signonsandiego.com/uniontrib/20060619/news_1s19usopen.html

Lewis, B., & Linder, D. (1997). Thinking about choking? Attentional processes and paradoxical performance. *Personality and Social Psychology Bulletin*, *23*, 937–944. doi: 10.1177/0146167297239003

Masters, R. S. (1992). Knowledge, knerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology*, *83*, 343–358.

Player, G. (2010). *Don't choke: A champion's guide to winning under pressure*. New York, NY: Skyhorse.

Savage, J., & Yancey, C. (2008). The effects of media violence exposure on criminal aggression: A meta-analysis. *Criminal Justice and Behavior*, *35*, 772–791.

Schell, M. J. (2005). *Baseball's all-time best sluggers: Adjusted batting performance from strikeouts to home runs*. Princeton, NJ: Princeton University Press.

Schlenker, B. R., Phillips, S. T., Boniecki, K. A., & Schlenker, D. R. (1995a). Championship pressures: Choking or triumphing in one's own territory? *Journal of Personality and Social Psychology*, *68*, 632–643. doi: 10.1037/0022–3514.68.4.632

Schlenker, B. R., Phillips, S. T., Boniecki, K. A., & Schlenker, D. R. (1995b). Where is the home choke? *Journal of Personality and Social Psychology*, *68*, 649–652. doi: 10.1037/0022–3514.68.4.649

Shipnuck, A. (2010). The no. 1 golfer of the year award: Phil Mickelson. *Golf Magazine*, *52*(10), 108–111.

Smith, A. M., Adler, C. H., Crews, D., Wharen, R. E., Laskowski, E. R., Barnes, K., … Kaufman, K. R. (2003). The 'yips' in golf: A continuum between a focal dystonia and choking. *Sports Medicine*, *33*, 13–31. doi: 10.2165/00007256–200333010-00002

Smith, A. M., Malo, S. A., Laskowski, E. R., Sabick, M., Cooney, W. P., III, … Kaufman, K. (2000). A multidisciplinary study of the 'yips' phenomenon in golf: An exploratory analysis. *Sports Medicine*, *30*, 423–437. doi: 10.2165/00007256–200030060-00004.

Wright, E. F., Jackson, W., Christle, S. D., McGuire, G. R., & Wright, R. D. (1991). The home-course disadvantage in golf championships: Further evidence for the undermining effect of supportive audiences on performance under pressure. *Journal of Sport Behavior*, *14*, 51–60.

Wright, E. F., Voyer, D., Wright, R. D., & Roney, C. (1995). Supporting audiences and performance under pressure: The home-ice disadvantage in hockey championships. *Journal of Sport Behavior*, *18*, 21–28.