Bulletin of the Psychonomic Society
1991, 29 (5), 384-386

"Bouncing back" from a loss: A statistical artifact

THOMAS R. ZENTALL
University of Kentucky, Lexington, Kentucky

Baseball teams that have won the World Series show the ability to bounce back after losing a game during the Series (Nahinsky, 1991). These World Champion teams are more likely to win a game if they have lost the previous game in the Series than if they have won the previous game. This apparent motivational effect is produced by a statistical artifact resulting from the constraint that ultimately they are the winner of the Series. On the other hand, for teams that eventually win the World Series, the probability of winning any individual game is greater than one would expect by chance, even if one corrects for the fact that they will be the Series winner.

The concept of motivation has played an important role in major theories of behavior (see, e.g., Hull, 1943). One of the factors that can contribute to motivational level is the extent of prior experience with success and failure. Furthermore, there is evidence that prior experience may affect not only one's *expectations* but also the actual *likelihood* of future success (i.e., an expectation can become a self-fulfilling prophesy; Jones, 1977).

Thus, one might not be surprised to learn that the probability that a baseball team wins a game in the World Series is affected by whether that team has won or lost the preceding game (Nahinsky, 1991). According to Nahinsky, however, and contrary to the intuition that baseball teams play in streaks or develop "momentum," a team that eventually wins the Series is more likely to win a game if it has *lost* the preceding game than if it has *won* that game. In other words, according to Nahinsky, the best baseball team "may be characterized by resilience," enabling it to be more likely to win the next game following a loss because of an increased level of motivation.

Nahinsky (1991) tracked the performance of winners of the 86 World Series played between 1903 and 1989 to arrive at his conclusion. He found that for Series winners, on the average, the probability of a win following a loss (.78) was significantly higher than the probability of a win following a win (.62).

Unfortunately, Nahinsky's calculations include a bias that ensures the outcome he found, independent of any sequential dependence between games. By considering only the eventual winner in assessing the conditional probabilities, Nahinsky introduced a necessary intergame dependency. Given a loss, there are necessarily fewer opportunities for a team to lose a game and still win the Series than there are given a win.

A simplified example will serve to illustrate this point. Let us assume that the Series involves the best two out of three games (rather than the normal four of seven). Now, if the eventual winning team has lost the first game, the probability that this team will win the second (and also the third) game must be 1.0 (otherwise, that team could not be the winning team). On the other hand, if that team had won the first game, it could lose the second game and still win the Series by winning the third game. Thus, in this case, and also in the case of a longer series, it must be that a winning team has a higher probability of winning following a loss than following a win.

To further demonstrate the bias produced by selecting only those teams that have won the World Series, one could actually identify all the possible sequences of game outcomes for the winning team, and then, setting the probability of winning a game at 0.5, one could calculate the theoretical probability of each conditional probability reported by Nahinsky (1991). As it turns out, there are only 35 different sequences of wins and losses possible for the winning team. First, given what Nahinsky calls the *stopping* rule (i.e., the fact that the Series stops as soon as one team wins four games), the actual number of games played will range from four to seven. Second, given that calculations are made only for the winning team, the outcome of the last game must be a win. Thus, the actual number of games free to vary will range from three to six. If one lists the allowable sequences, one will find only 1 way to win in four games, 4 ways to win in five games, 10 ways to win in six games, and 20 ways to win in seven games (see Table 1). Next, to calculate the probability of a win given a win in the previous game, $P(W2 \mid W1)$, for each pair of games, one would count the number of sequences in which both games of the pair are won and divide that number by the total number of sequences in which the first game of the pair is won. Similarly, to calculate the probability of a win given a loss in the previous game, $P(W2 \mid L1)$, one would count the number of sequences in which the first game is lost and the second

The author thanks Janice Steirn for her helpful comments. Correspondence should be sent to Thomas R. Zentall, Department of Psychology, University of Kentucky, Lexington, KY 40506.

Table 1
Possible Sequences of Wins and Losses for the Winning World Series Team

Game 1	Game 2	Game 3	Game 4	Game 5	Game 6	Game 7
		Four-Game Series (1)				
W	W	W	W			
		Five-Game Series (4)				
L	W	W	W	W		
W	L	W	W	W		
W	W	L	W	W		
W	W	W	L	W		
		Six-Game Series (10)				
L	L	W	W	W	W	
L	W	L	W	W	W	
L	W	W	L	W	W	
L	W	W	W	L	W	
W	L	L	W	W	W	
W	L	W	L	W	W	
W	L	W	W	L	W	
W	W	L	L	W	W	
W	W	L	W	L	W	
W	W	W	L	L	W	
		Seven-Game Series (20)				
L	L	L	W	W	W	W
L	L	W	L	W	W	W
L	L	W	W	L	W	W
L	L	W	W	W	L	W
L	W	L	L	W	W	W
L	W	L	W	L	W	W
L	W	L	W	W	L	W
L	W	W	L	L	W	W
L	W	W	L	W	L	W
L	W	W	W	L	L	W
W	L	L	L	W	W	W
W	L	L	W	L	W	W
W	L	L	W	W	L	W
W	L	W	L	L	W	W
W	L	W	L	W	L	W
W	L	W	W	L	L	W
W	W	L	L	L	W	W
W	W	L	L	W	L	W
W	W	L	W	L	L	W
W	W	W	L	L	L	W

Note—Values in parentheses represent the number of ways there are to win, given the respective number of games (1 way to win in a four-game series, etc.).

game is won and divide that number by the total number of sequences in which the first game is lost.

If one carries out these calculations for each successive pair of games (omitting the last game of each sequence), one will obtain the conditional probabilities listed in Table 2. The observed conditional probabilities associated with the World Series data reported by Nahinsky (1991) are also presented. For comparison, the differences in conditional probabilities for both calculated and observed data are presented in Table 2, as well.

The comparison that most pertains to the hypothesis in question (viz., that the Series winner is more motivated to win following a loss than following a win) is the difference between $P(W2 \mid L1)$ and $P(W2 \mid W1)$ for both random and observed sequences. As is indicated in Table 2, the difference in *calculated* conditional probabilities pooled over the various pairs of games (.18) is quite comparable to the difference in *observed* conditional probabilities (.16). Thus, on the average, the difference in conditional probabilities is not different from what one would expect by chance, given independence between games.

It is interesting that the major discrepancies from chance reported by Nahinsky (1991) occurred in Games 3 and 6. Given that the first two and last two games in a seven-game series are generally played on the home field of one of the teams, whereas the middle three games are played on the home field of the other team, is it possible that the major discrepancies from chance occur when an (eventual) Series winner returns home after a loss? On the other hand, if this effect occurs, it is counteracted by the somewhat greater tendency than one would expect by chance to win following a win (rather than a loss) in the other games (especially in the fourth game).

Examination of Table 2 also indicates that the calculated conditional probabilities themselves are consistently lower than the observed conditional probabilities. Given that only allowable sequences were considered and performance on the final game was not included (because it was not free to vary), these findings suggest that the assumption of equal probability of a win or a loss by the eventual winner of the Series was not correct. In other

Table 2
Calculated Probabilities of Winning a World Series Game, Given that the Previous Game Was Won [$P(W2 \mid W1)$] or Lost [$P(W2 \mid L1)$], Versus the Observed Probabilities

	Calculated			Observed*		
Game	$P(W \mid W)$	$P(W \mid L)$	Difference	$P(W \mid W)$	$P(W \mid L)$	Difference
2	.50	.67	.17	.59	.71	.12
3	.50	.67	.17	.53	.90	.37
4	.47	.67	.20	.75	.66	−.09
5	.44	.64	.20	.67	.79	.12
6	.40	.60	.20	.56	.86	.30
Total	.47	.65	.18	.62	.78	.16

*From Nahinsky (1991) Table 1.

words, given that a team has won the Series, the probability that it will have won any single game (not including the last game) is actually greater than .5. Thus, these results suggest the comforting conclusion that in past World Series competition, one of the teams may have been better, on the average, than the other. For an eventually winning team, given that the prior game was lost, the difference between the observed probability of winning the next game and the calculated probability of that event was .13. Similarly, for such a team, given that the prior game was won, the difference between the observed probability of winning the next game and the calculated probability of that event was .15. Thus, one can conclude that the outcome of the World Series is not a completely random, equiprobable event.

REFERENCES

HULL, C. L. (1943). *Principles of behavior: An introduction to behavior theory*. New York: Appleton-Century-Crofts.

JONES, R. A. (1977). *Self-fulfilling prophecies*. Hillsdale, NJ: Erlbaum.

NAHINSKY, I. D. (1991). Bouncing back in the World Series. *Bulletin of the Psychonomic Society*, **29**, 131-132.

(Manuscript received March 15, 1991.)