Scand J Med Sci Sports 2012: 22: e64–e69
doi: 10.1111/j.1600-0838.2011.01408.x



SCANDINAVIAN JOURNAL OF
MEDICINE & SCIENCE
IN SPORTS

# Accuracy of professional sports drafts in predicting career potential

**D. Koz, J. Fraser-Thomas, J. Baker**

*School of Kinesiology and Health Science, York University, Toronto, Canada*
*Corresponding author: J. Baker, School of Kinesiology and Health Science York University, 4700 Keele St., Toronto, ON, Canada M3J 1P3. Tel: +1 416 736 2100 ex. 22361, Fax: +1 416 736 5774, E-mail: bakerj@yorku.ca*

Accepted for publication 31 August 2011

**The forecasting of talented players is a crucial aspect of building a successful sports franchise and professional sports invest significant resources in making player choices in sport drafts. The current study examined the relationship between career performance (i.e. games played) and draft round for the National Football League, National Hockey League, National Basketball League, and Major League Baseball for players drafted from 1980 to 1989 ($n$ = 4874) against the assumption of a linear relationship between performance and draft round (i.e. that players with the most potential will be selected before players of lower potential). A two-step analysis revealed significant differences in games played across draft rounds (step 1) and a significant negative relationship between draft round and games played (step 2); however, the amount of variance accounted for was relatively low (less than 17%). Results highlight the challenges of accurately evaluating amateur talent.**

The professional sports draft is a process used among all major sports in North America to distribute players among a pool of eligible teams. The purpose of the draft is to give exclusive rights to a team to sign a player to a contract. The best-known type of draft is the entry draft, used to allocate players who have recently become eligible to play in a league. Varying slightly from sport to sport, players can come from junior teams, high schools, or colleges from around the world. Theoretically, the entry draft deters costly bidding wars for talented young athletes and ensures that no team can monopolize all the best young players, thereby making the league unbalanced (Popper, 2004).

There is considerable economic risk with the drafting process, as the first pick in the draft can sign an initial contract up to four times the amount of the last pick in the first round (Massey & Thaler, 2010). This implies that a player drafted earlier will be much more successful and will bring more market value to their team than players drafted later (i.e. either later in the same round or in a later round). Previous research has shown, however, that managers of professional teams significantly overvalue early draft choices over late draft choices without solid economic rationale (Massey & Thaler, 2010).

Staw and Hoang (1995) studied draft order in the National Basketball Association (NBA) to determine the relationship between the round in which a player was drafted and the playing time that player received. The researchers examined this association from an economics framework and focused on the concept of "sunk costs" (i.e. costs that cannot be recovered once they have been incurred such as drafted players' initial salaries). What Staw and Hoang discovered was that highly paid early round draft picks were not statistically superior players than those drafted in later rounds. Overall, early round draft picks had longer survival rates only because of the costs incurred upon them after they were drafted. More specifically, draft order was a significant predictor of survival rate in the NBA after controlling for on-court performance.

In another study examining draft order and performance, Berri and Simmons (2009) focused on the performance of quarterbacks in the National Football League (NFL) relative to three main research questions: (a) what is the relationship between an NFL quarterback's draft position and his subsequent performance?; (b) what factors do NFL teams consider in drafting a quarterback?; and (c) how do these factors relate to subsequent performance? The relationship between actual performance and draft position was weak; interestingly, they found a stronger relationship between total plays and draft position, suggesting that draft position gets you on the field more, but quarterbacks who are chosen earlier in the draft do not appear to outperform those chosen later.

Most recently, Boulier et al. (2010) examined the relationship between draft round and performance among quarterbacks and wide receivers in the NFL. They con-

sidered three measures of performance for quarterbacks (i.e. years played, number of passes thrown, and quarterback rating), as well as two measures of performance for wide receivers (i.e. number of years in the league and total receiving yards). Their analysis revealed that the earlier a quarterback was drafted, the better his performance. More specifically, earlier drafted quarterbacks had a higher quarterback rating, threw more passes, and played for more years. These findings reinforce previous work, in that Berri and Simmons showed that earlier drafted players received more playing time, and therefore, also had higher attempts and games played. However, their findings conflict with Berri and Simmons showing that earlier draft pick was associated with higher quarterback rating relating to higher performance over a career. Wide receivers showed a similar trend; those drafted earlier played for more years and had more receiving yards. The researchers concluded that football executives made very good draft picks.

Hypothetically, the most effective process in the draft would be to select the most desirable player in round one, followed by the second most desirable player in round two, and so on (i.e. a negative linear relationship) based on the assumption that as draft round increases, player potential decreases. Although there is some support for this assumption in the work summarized above, it is based on very limited evidence and there is considerable room for discovery regarding the accuracy of professional scouts and general managers at selecting the best athletes. There have been few examinations of either the National Hockey League (NHL) or the Major League Baseball (MLB) and data from the NFL have been limited to quarterbacks and wide receivers. As such, this study builds on previous research by focusing on the four main professional sports in North America. Specifically, the purpose of this study was to determine whether draft order predicts athletic potential in the NBA, NFL, NHL, and MLB.

## Methods

### Participants

To maintain consistency across sports, the sample included players drafted into the NHL, NFL, NBA, and MLB from 1980 to 1989 who recorded one or more games played. One thousand fifty-nine players from MLB, 407 from the NBA, 2380 from the NFL, and 1028 from the NHL made up the resulting sample. Given that most players drafted during this time period had finished their careers at the time of data collection, complete career statistics were collected in most cases. The number of individuals who had not completed their career was minimal (less than 1% of players) and was not expected to have an effect on the analysis. For the NHL and NFL, the entire sample of drafted players was used. For the NBA, only the first two rounds of the draft data were used. This was done to maintain consistency with previous work (Staw & Hoang, 1995) as well as to be applicable to the current draft structure. For MLB, a different approach was necessary, given that MLB has two main drafts and two supplementary drafts each year. The MLB sample in the current analysis was limited to the first 15 rounds of the June amateur draft only for two reasons. First, the June amateur draft is the primary entry draft for this league. Second, we wanted to maintain relative consistency with the other leagues in this study in the number of rounds examined.

### Data collection

The performance data for the sample populations were collected through the official online resources of the professional leagues (i.e. http://www.nhl.com, http://www.nfl.com, http://www.mlb.com, and http://www.nba.com). These websites contain information on each year of the corresponding sports' drafts since the leagues' inceptions, as well as links to individual players' statistics. Specifically, they are the primary sources of data on players in the professional leagues as they house information on each player who has been drafted and has competed at the professional level of competition (i.e. has a statistic recorded in the official database). Reliability of the data collected was determined through a cross examination of a random selection of 10% of the athletes' statistics with two official encyclopedias of North American professional sports (Shouler, 2004; Gaschnitz, 2008). This reliability check resulted in perfect (i.e. 100%) consistency between the online sources and the published encyclopedias.

### Measures

The databases provide a considerable range of player performance indicators. In an effort to maintain consistency across the sports and positions, *games played* was used as the standard outcome variable as it is the only variable that crosses all sports and positions. Moreover, a preliminary analysis (Koz et al., 2009) showed that games played accounts for 80% of the variance in the points scored by NHL players drafted from 1980 to 1989, increasing the construct validity of this variable as an appropriate measure of performance. Further, it can be argued that if players are accumulating playing time, then they must be contributing sufficiently to be considered an asset to their franchises. More simply, an athlete's value to a team (i.e. potential) is reflected in coaching decisions to play the athlete as often as possible. The predictor variable in the current analyses was *draft round*.

### Data analysis

The basis for the analyses outlined below is the hypothetical negative linear relationship between player potential and draft round noted previously. Inherent in the draft structure (i.e. draft order determined by the previous season's outcome, with the teams with the worst records selecting before teams with the best records) is the assumption that player potential differs for the incoming athletes, and a team's draft position matters. More specifically, there should be a relationship evident in the overall production of the players selected in earlier vs later rounds, or draft round and selection order would serve little purpose. If incoming players had the same potential for contribution to the teams in the league, the structure could be a free-for-all selection of a pool of players until the teams had satisfied a certain player quota for their next season's training camp. Instead, the draft structure allows an orderly selection of the incoming talent into the league.

To examine whether professional drafts are accurate at predicting player potential, we used a two-step approach. First, we determined whether there was a difference in games played across draft rounds using Kruskal–Wallis (K–W) tests (effect size is reported as $\phi$). We ran five separate analyses: one for NBA, NHL, and NFL, and two for MLB, separating the baseball data into pitchers and fielding players. Independent analyses were conducted for pitchers and fielding players because pitchers work on a rotation basis, usually with 5–6 regular starting pitchers per franchise, whereas fielding players (barring any injuries) usually play in most of the

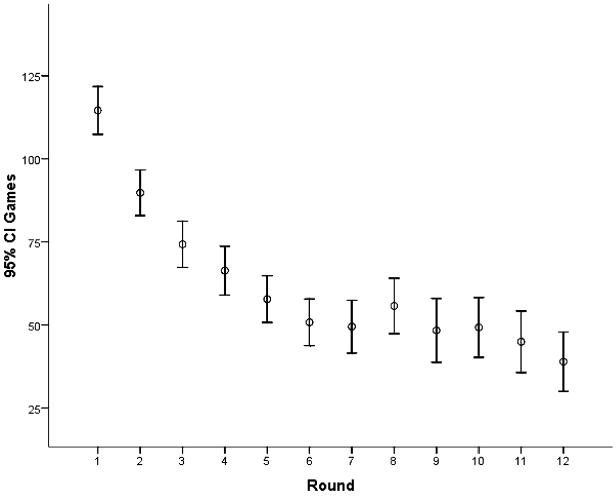


*Fig. 1.* NFL mean games played across draft round.

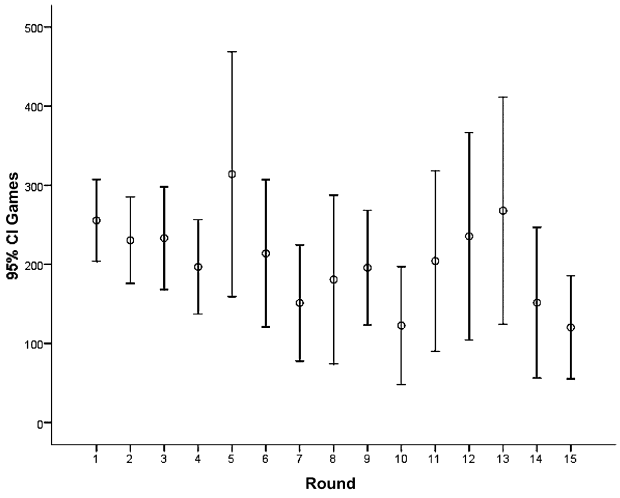


*Fig. 2.* MLB mean games played for pitchers across draft round.

regular season games. Test variables included games played and draft round, with the corresponding ranges based on the number of draft rounds in the league being analyzed (i.e. 1–12 for NFL, 1–15 for the MLB, 1–10 for the NHL, and 1–2 for the NBA). Although nonparametric tests of difference (i.e. K–W test) were most appropriate for the data, post-hoc analyses can be difficult to administer and interpret. Therefore, we ran parallel analyses with analysis of variance and where appropriate (i.e. when parametric and nonparametric analyses were in agreement), post-hoc tests were conducted using Tukey's method. The second step in the analysis involved a hierarchical regression analysis with games played as the dependent variable. In level 1, position was entered followed by draft round in level 2. Although the estimate of effect size may be conservative (compared to a simple correlation between games played and draft round), it removes any possible confounding influence of differences between positions. Each league was analyzed separately, with two separate analyses again for MLB pitchers and fielders. All analyses were performed using SPSS version 17.0 (SPSS Inc., Chicago, Illinois, USA) and statistical significance was set at $P < 0.05$.

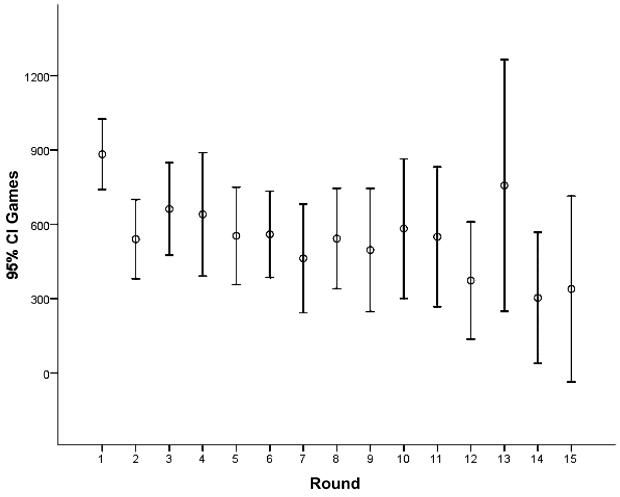


*Fig. 3.* MLB means games played for fielding players across draft round.

## Results

### NFL

Figure 1 shows the mean number of games played across draft round. Results from the K–W analysis indicated small- to medium-sized statistically significant effect for number of games played across the draft rounds [$\chi^2$ (11) 377.16, $P < 0.05$, $\phi = 0.34$]. Post-hoc analyses indicated that round 1 differed from all other rounds. Round 2 differed from all other rounds except round 3. Round 3 differed from rounds 2, 4, and 5. Round 4 differed from rounds 1, 2, and 11. Round 5 only differed from rounds 1 and 2. Rounds 6 to 10 differed from rounds 1, 2, and 3. Round 11 differed from rounds 1, 2, and 4 while round 12 differed from rounds 1 to 4. Results of the regression analysis revealed a significant effect of position on games played in level 1 ($\beta = 0.10$, $P < 0.001$) followed by a significant negative relationship for draft round ($\beta = -0.32$, $P < 0.001$) in level 2.

### MLB

Figures 2 and 3 present the mean number of games played by draft round for MLB pitchers and fielders, respectively. Results from the K–W analysis for MLB pitchers did not show a significant difference in the number of games played across the draft rounds [$\chi^2$ (14) 23.354, $P = 0.055$, $\phi = 0.12$] although the test was significant for fielding players [$\chi^2$ (14) 33.482, $P < 0.05$, $\phi = 0.13$]; however, both effect sizes were low. Post-hoc analysis of this relationship in the fielding players indicated that while no rounds were significant from each other at the $P < 0.05$ level, the difference between round 1 and rounds 2 and 7 approached significance ($P < 0.10$). Unlike the other analyses, the regression analysis for pitchers only had one step (i.e. because there was only one position) and there was a significant negative effect

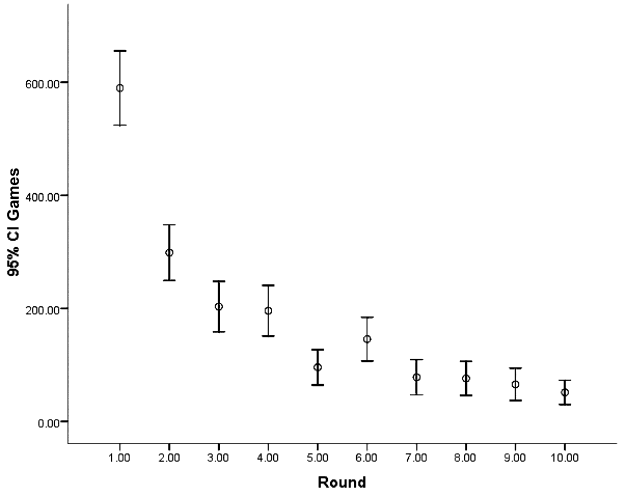


*Fig. 4.* NHL mean games played across draft round.

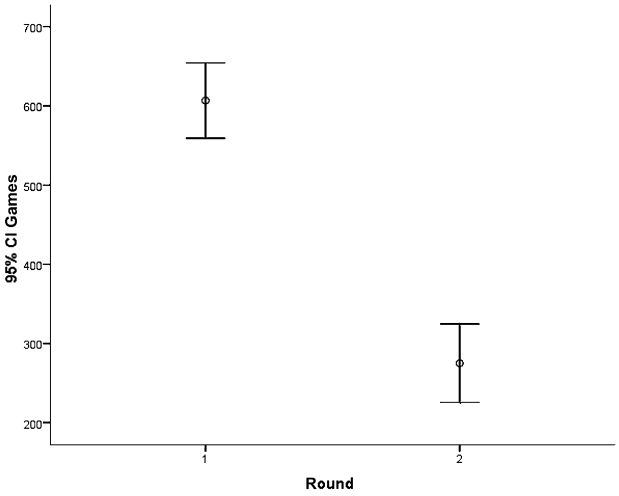


*Fig. 5.* NBA mean games played across draft round.

for draft round ($\beta = -0.11$, $P < 0.05$). For fielding players, there was a significant negative relationship for draft round on games played ($\beta = -0.16$, $P < 0.001$) but no effect for position ($P = 0.97$).

### NHL

The mean number of games played across draft round for the NHL is presented in Fig. 4. Results indicated a medium-sized significant effect for the number of games played across the draft rounds [$\chi^2$ (9) 519.559, $P < 0.05$, $\phi = 0.50$]. Post-hoc analyses indicated that rounds 1 and 2 differed from all other rounds, while round 3 differed from all rounds but rounds 4 and 6. Round 4 differed from rounds 3 and 6. Round 5 differed from rounds 6 to 10. Round 6 differed from rounds 3 to 9, and round 7 differed from rounds 5 to 10. Rounds 8 and 9 both differed significantly from rounds 5 to 10. Finally, round 10 differed from rounds 5 and 7 to 10. The regression analysis indicated a significant negative relationship for draft round on games played in the NHL ($\beta = -0.37$, $P < 0.001$) even after controlling for the significant effect of position ($\beta = 0.08$, $P < 0.01$).

### NBA

Figure 5 presents the mean number of NBA games played across draft round for the NBA. Results from the K–W analysis indicate a medium-sized statistically significant difference between the distributions of games played across the two draft rounds [$\beta^2$ (1) 85.309, $P < 0.05$, $\phi = 0.46$]. In level 1 of the regression analysis, the relationship between playing position and games played was not significant ($P = 0.94$); however, there was a moderate-sized significant negative relationship for draft round in level 2 ($\beta = -0.42$, $P < 0.001$).

## Discussion

Consistent with past works (Staw & Hoang, 1995; Boulier et al., 2010), there was a negative relationship between draft round and games played across the leagues indicating that players drafted earlier were used more by their teams than those drafted later. Players drafted into the NBA in the second round played less than those drafted in the first round, although with only two rounds in the NBA draft, it is difficult to conclude too much from this league alone. There were numerous differences in games played across draft rounds in every draft but MLB, most notably round 1, differed significantly from all other rounds with some intermediate differences in the draft rounds of the NHL and NFL. It is possible that players drafted in round 1 are the most talented among the incoming athletes and as a result have the most games played, but as prior research has noted (Staw & Hoang, 1995), these players are most affected by the "intangible elements" that can confound the draft (sunk cost, escalation of commitment). While our analyses only indicate that this group is generally unique compared to other draft rounds in terms of their games played, future studies of these early round draftees is necessary to determine the legitimacy of their exceptional "potential."

The consistency in the pattern of findings may provide some evidence of the efficacy of the professional draft at identifying talent. It is noteworthy, however, that even the strongest relationship identified in the regression analyses accounted for less than 17% of the variance in games played leaving a considerable amount of variability between players explained by other factors (e.g. injury, late "blooming," etc.). Moreover, when important confounds such as "sunk costs" are added to this discussion, the ability of the draft to identify talent seems low.

The results for MLB are intriguing. There were no significant differences between the draft rounds for

games played (although the overall effect was significant for fielders) and the size of the association between these variables was considerably smaller than in the NFL, NHL, and NBA. Initially, this seemed contrary to prior work. For instance, Spurr (2000) found that draft position predicted the chance of logging a game played in the major leagues but these differences may be due to different outcomes across the two studies. More specifically, we considered total games played while Spurr (2000) used a binary outcome determined by whether the player played a single game in the league. The former is a more accurate measure of a player's potential, as total games played is a better representation of the value of a drafted player.

The data for MLB pitchers were particularly unique compared to the other leagues. This may be due to the different talent development system professional baseball has compared to the other leagues. Almost all players who are drafted into MLB go through lengthy careers in the minor leagues and this prolonged "trial" period before entering the major leagues may be a better indicator of their ability to compete at the professional level. A player's performance in the minor league system over time could wash out any high draft selection biases that affect other leagues, such as the sunk cost effect. While this explanation seems reasonable, it is likely that other factors contribute to this effect since the NHL also has a "farm system" where drafted players hone their skills before being called up to the professional ranks, yet the NHL shows a similar trend to the NFL and NBA. A more in-depth analysis of MLB, both individually and in comparison to the other leagues, would expand our understanding of the factors affecting games played at the highest level of this professional sport.

Although this study provides a useful step in developing our understanding of professional sports drafts, there were some limitations to our analysis. First, this study could not control for injuries sustained by players, which can shorten their career prematurely. However, players drafted from every round are exposed to injury risk and the large sample of players from each of the four leagues should have minimized the overall influence of this confounding variable, but this should be verified with future work.

Second, our examination was limited to the data provided by the official archives, which limited our analyses in some instance. For example, pitchers in MLB can be either starting or relief pitchers and because relief pitchers play fewer innings, they can play in more games than starters who work on a rotation (usually play every five games). Our data did not include this measure (and this designation typically changes across a pitcher's career); however, researchers may wish to control for this influence in future analysis.

Finally, the timeframe of 1980–1989 for drafted players was chosen because nearly all of these players had completed their careers thereby allowing access to career statistics. However, this may limit the current relevance of our analyses since drafting techniques may have changed over the past 20 years. Although it is unlikely that these changes have been so drastic as to completely change the pattern of results, future studies are necessary to determine the longevity of the patterns seen in these data. The NBA may be an exception since its draft process underwent significant changes in the 1980s. In 1985, the draft lottery was introduced, and in 1988, the number of draft rounds decreased from seven to three, and then in 1989, it changed to the current structure of two rounds. Only the two rounds of data were included in our analyses because this is most applicable to the modern structure, but it is possible that drafting strategies differed in the early and mid-1980s when more rounds of the draft were available.

This study highlights several areas for future research. For instance, adding performance statistics and/or collecting minor league performance data may add further depth to our understanding of talent development after the draft and explain the discrepant results for MLB. Additionally, it may be important to determine which teams are better or worse at accurately drafting talented players. Any significant differences in a team's performance in the draft over time may inform the development of more effective drafting strategies. Most significantly, we have only a rudimentary understanding of the possible mechanisms driving these effects. While the "sunk cost effect" is a plausible mechanism, there are others; for instance, reputations of the scouting and coaching staff are largely based on their ability to recognize and develop players with the greatest potential, which may perpetuate a system where players with similar capabilities but drafted earlier get greater opportunities than those drafted later simply to show that scouting and coaching decisions were accurate.

Players drafted into professional sports often come with high expectations from media, fans, and teammates. Whether these players meet those expectations or fulfill their potential is the result of a dynamic interaction among a host of variables including player characteristics, team environment, and social pressure. Our analysis showed similar trends between games played and draft round in the NFL, NBA, and NHL with MLB as the anomaly. These findings provide a foundation for future analyses to inform our understanding of the drafting process and improve the accuracy of professional franchises in predicting potential.

## Perspectives

While increasing in importance in both professional and amateur sport systems, accurate identification of athlete potential (i.e. talent) remains elusive. Moreover, the processes used to identify and nurture talented athletes are confounded by subjective biases (e.g. "sunk costs"). In the present study, there was a consistent negative

relationship between draft round and games played but the size of these effects were generally small and varied by sport. Interestingly, the baseball draft seemed particularly ineffective compared to other sports, although this may reflect differences in player development systems across sports. In sum, our results highlight the limited accuracy of the current draft system in professional sports and emphasize the considerable room for improvement in this area.

**Key words:** performance, talent identification.

## References

Berri D, Simmons R. Catching a draft: on the process of selecting quarter-backs in the National Football League amateur draft. J Prod Anal 2009: 35: 37–49. [cited 12 January 2010]. Available from http://www.springerlink.com/content/k96t8116v8686350/fulltext.html

Boulier BL, Stekler HO, Coburn J, Rankins T. Evaluating National Football League draft choices: the passing game. Int J Forecast 2010: 26: 589–605.

Gaschnitz M. Statistical encyclopedia of North American professional sports. New York: McFarland & Co, 2008.

Koz DR, Baker J, Fraser-Thomas J. NHL draft is a poor predictor of athletic potential: preliminary results. Poster session presented at the annual meeting of The Canadian Society for Psychomotor Learning and Sport Psychology (SCAPPS), Toronto, ON, 2009.

Massey C, Thaler RH. The losers curse: overconfidence vs market efficiency in the National Football League draft. University of Chicago, mimeo, 2010. [cited 13 September 2010]. Available from SSRN: http://ssrn.com/abstract=697121

Popper S. Draft strategy: how has it changed? Athlon Sports Pro Basketball 2004: 11: 19–21.

Shouler K, ed. Total basketball: the ultimate basketball encyclopedia. Toronto, ON: Sports Media, 2004.

Spurr S. The baseball draft: a study of the ability to find talent. J Sports Econ 2000: 1: 66–85.

Staw MB, Hoang H. Sunk costs in the NBA: why draft order affects playing time and survival in professional basketball. Adm Sci Q 1995: 40: 474–494.