Psychology of Sport and Exercise 16 (2015) 128–136

Contents lists available at ScienceDirect

Psychology of Sport and Exercise

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

Predicting success in the National Basketball Association: Stability & potential

Jerad H. Moxley*, Tyler J. Towne

Florida State University, United States

A R T I C L E I N F O

Article history:
Available online 22 July 2014

Keywords:
Basketball
Skill
Anthropometrics
Talent
Forecasting
Matthew effect

A B S T R A C T

Objectives: To create a more rigorous model of early career success among players in the National Basketball Association (NBA) using growth mixture models. To test the extent to which NBA careers can be predicted by variables that represent past performance and variables that might represent untapped potential.
Design: Archival data was collected from measures taken at the pre-draft NBA combine and publicly available data on college and NBA performance.
Method: The first three years of players' NBA careers from 2001 to 2006 draft classes were analyzed using a growth mixture model with collected variables predicting latent class. The estimated parameters were then used to forecast the 2007 to 2010 draft classes. Draft order was also predicted with the same variables.
Results: NBA player skill formed 3 latent classes of players; only one class performed well in the NBA. Membership in the strongest class was only predicted by age, quality of college program, and players' college performance. Latent class probabilities predicted NBA career trajectory slightly better than draft order in both the estimation model and in the forecast model. NBA draft order was predicted by the same variables as well as arm span and agility.
Conclusions: None of the variables analyzed supported an “untapped potential” hypothesis. There is clear evidence for roles of training environment and the stability of skill. The data is consistent with views of deliberate practice and skill acquisition and appears to be consistent with data showing the benefits of being identified as talented, such as the Matthew effect.

In professional team sports, where a large amount of money is at stake, identifying players who are the most likely to contribute to a team's future success is important. This is particularly true during the transition to a higher level of competition such as from college to professional sports. During this time, the competition improves and scouts and general managers attempt to use information about players' past performance and attempt to measure their future potential. There is a high financial incentive for teams to identify players with the potential for success beyond what is indicated by their past performance (Berri & Schmidt, 2006). The focus of this paper is to examine factors that are most likely to contribute to early career success among professional players in the National Basketball Association (NBA) and to better understand the factors available on draft day that significantly predict players' future performance.

Before and during the draft, teams' primary objective is to identify players who will perform at a high level and make immediate contributions. Selecting the right player in a draft can increase the value of a franchise by millions of dollars (Hausman & Leonard, 1997). The opportunity to make an early draft choice is so valuable that NBA teams appear to reduce their effort to win games once they have been eliminated from the playoffs, thereby improving their probable draft position (Walters & Williams, 2012).

During the draft process, teams' general managers, coaches, and scouts examine many variables. Chief among these variables is players' previous performance. Teams are particularly interested in players who have distinguished themselves at lower levels of competition such as high school, college, club, semi-professional, or other professional leagues. In addition to this obvious marker of skill, there are other factors that are taken seriously such as players' height, arm span, vertical leap, hand size, and agility. To develop clear hypotheses about which factors are most predictive of future performance, it is important to understand general observations about how skill develops across domains.

* Corresponding author.
E-mail address: moxley@psy.fsu.edu (J.H. Moxley).

http://dx.doi.org/10.1016/j.psychsport.2014.07.003
1469-0292/

Searching for talent

When studying skill, and particularly when forecasting changes in skill, terms are often theoretically loaded. For example, talent can be viewed as one's current level of skill or, perhaps, how skilled one could be given optimal training conditions. Within the context of this paper we will define “talent” as a player's potential ceiling of performance. More specifically, we will not define talent by previous performance alone. It is worth noting that talent accounts also argue that talent may not represent different ceilings of performance but rather the ability to learn more quickly, or some combination of the two. For the sake of simplicity, we will simply define talent as different ceilings of performance.

Prior to the NBA draft, talent should have partly manifested itself through previous performance (e.g. there is no evidence to our knowledge of players with no previous basketball experience being drafted). Therefore, we will use “untapped potential” to refer to the idea that two equally skilled individuals, could be different distances from the limits of their talent. When scouts are examining players for the NBA draft, they are attempting to not only identify those with proven performance but also players with greater untapped potential. We will use the term “skill” to refer to a player's current level of performance. In sum, talent can be conceptualized as containing two components, current skill level plus untapped potential. With this conceptualization, scouts and managers in the NBA look to identify individual differences in talent (value) by taking into account current skill levels and identifying factors that they believe are representative of untapped potential.

Each year NBA teams are looking for untapped potential when they discuss their draft pick. The term “upside” is often used to refer to a player's ceiling and is used commonly to defend decisions. For example, when discussing the surprising number one draft pick of Anthony Bennett in 2013, Cleveland Coach Mike Brown stated, “He's got long arms … he has a lot of upside” (Finnan, 2013). Even when discussing 2012 number one draft pick Anthony Davis who, as a freshman, was the consensus player of the year, helped lead his team to the national championship, and lead the NCAA in win shares, Hornets General Manager Dell Demps noted, “his offensive upside is great” (Associated Press, 2012).

One of the complications of research on superior performance is that it is difficult to partial out talent from various environmental influences. For example, several studies have demonstrated the “season-of-birth effect” or “relative age effect” which shows that individuals are born before the age cut-off in their selected sports are more likely to reach professional levels, ostensibly because they happen to be older and therefore more physically and cognitively developed, on average, than others in their cohorts. Coaches identify these players as being more talented and offer them more opportunities to improve their skill. This effect has been identified in hockey, soccer, baseball, gymnastics, and other sports (see Musch & Grondin, 2001 for a review). Athletes identified as talented also tend to have access to the best coaches and trainers, as well as parents or guardians who are committed financially and personally to their success (Bloom, 1985). Similarly, Merton (1968) summarized a wide body of evidence demonstrating Matthew effects such that, even unearned, opportunities and positions could generate lasting benefits to a person's professional success.

One of the clearest sources of skill is deliberate practice within the domain. Studies have consistently found that world-class performers engage in massive amounts of practice in many cases these individuals practice 4–5 h every day for several years (Charness, Tuffiash, Krampe, Reingold, & Vasyukova, 2005; Ericsson, Krampe, & Tesch-Römer, 1993). In studies across domains, between 40 and 60% of the variance in skill could be explained by deliberate practice history (Charness et al., 2005; Ericsson et al., 1993). When identifying skill, one must primarily look to the amount of practice and its associated competitive outcomes to pinpoint those individuals who are most likely to excel.

Finally, the idea that there are individual differences in the limits of skill acquisition has been popular in the academic and sports literature, at least since Sir Francis Galton performed the first systematic studies investigating such possibilities. By examining families of eminent individuals and measuring simple characteristics such as reaction time, Galton concluded that those who were particularly talented could achieve much higher levels of performance than their less talented counterparts (Galton, 1869/1979). The idea of this potential based on measurable basic abilities persists to this day, and seems to guide many of the decisions made by NBA teams when drafting players.

Based on this information, scouts may either rely on previous performance, which we will argue is consistently related to future performance given reasonable temporal proximity, or on perceived untapped potential; factors that are assumed to uniquely relate to their ability to adapt and prosper at the next level of competition. For scouts, it is important to note that a variable is only interesting if it is systematic and predictable by an observable metric. Fig. 1 shows a theoretical model that we believe guides researchers and general managers when forecasting NBA talent. Ericsson et al. 1993 showed that skill progresses through a series of power functions with plateaus giving way to increases in performance at transitional periods. Deliberate practice accounts propose that the individual differences in improvement rates after transitions are primarily due to practice activities and access to improved training resources. Untapped potential accounts would attribute at least part of any discovered individual differences in improvement rates to athletic characteristics of the type measured during the combine.

Stability of performance

The development of reliably superior skill takes place across years or even decades of gradual improvement (Ericsson et al., 1993). When skill is plotted against time it is clear that there are dramatic early increases over a relatively short amount of time, but as skill increases the gains become more and more slight. Therefore, when examining high-level players, performance at any two temporally close events should be highly correlated.

Chess gives us one of the clearest examples of this type of improvement. Chess is particularly attractive to skill researchers

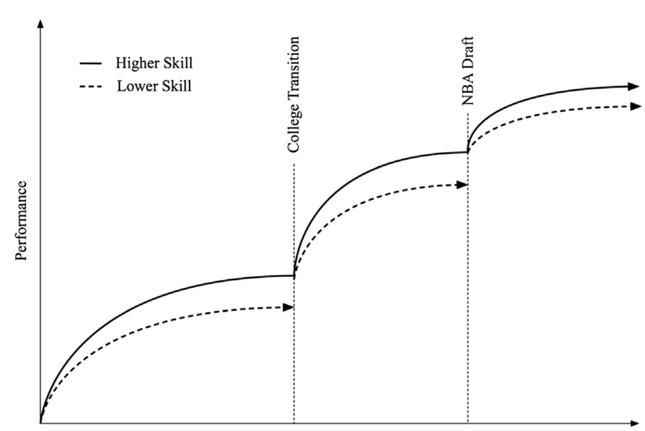

Fig. 1. Shows the trajectories of a group of hypothetical players. Only higher skilled players make transitions from one level of play to the next, at which point, the individual differences among them determine value to the team. According to the “untapped potential” hypothesis, part of the differences in skill at each level can be accounted for by athletic variables measured at the combine.

because it has one of the most well established rating systems of any domain. These ratings are computed based on previous tournament performance and have been shown to be a reliable way of predicting the outcome of a given tournament game (Elo, 1978). Improvement in chess, especially at high levels, is slow and gradual as is subsequent decline with age, and players' previous tournament performance explains most future tournament outcomes, with small negative effects from aging (Moxley & Charness, 2013).

In music, Ericsson, et al. (1993) showed that skill was strongly related to the amount of deliberate practice that the individual had engaged in. The most accomplished performers, those who would go on to be concert musicians, had dedicated over 5000 more hours to deliberate practice than those who would go on to be music teachers. Studies have suggested that skill increases monotonically with deliberate practice as a power function and the most skilled participants within a domain will overcome a series of plateaus, at which time performance must be adjusted to foster continued improvement and avoid automaticity. These plateaus are typically found during transitional periods, such as times when casual engagement gives way to serious involvement, or when a participant decides to engage in the domain full-time, as is the case when athletes transition to the professional level (see Fig. 1) (Ericsson et al., 1993). There is no credible evidence in any competitive domain that someone who has been identified based on aptitude or innate talent has reached world-class levels of performance without recording the prerequisite amount of deliberate practice (Ericsson, Roring, & Nandagopal, 2007).

However, it is possible that sports require unique qualities that allow more expression of untapped potential than in a domain such as music. For instance many sports have advantages for those of specific body sizes. It may be easier for someone who is 6′7″ to excel at basketball than it is someone who is 5′1″. According to the official NBA website, the average height of NBA players in the 2007–2008 season was 6′ 6.98″; well above the CDC reported national men's average of 5′ 9.3″ (Fryar, Gu, & Ogden, 2012). It is also likely that tall men are given more opportunities at early ages to engage in the sport. In this paper we attempt to determine whether performance in the NBA can be reliably predicted by physical metrics such as height and agility, but critically, when controlling for previous competition performance.

Quantifying athleticism

Much like in cognitive domains where memory and intelligence are sometimes measured in an attempt to assess talent, professional sports teams evaluate players based on basic athletic measures. For every player entering the 2013 draft the official NBA web page lists their height (with and without shoes), weight, arm span, standing reach, body fat, hand length, and hand width. Other measures of player athleticism are available to the public, including their vertical leap, bench press, "lane agility", and ¾ court sprint.

Psychological studies have sought to identify other basic features of athletes at key positions such as goalies in hockey and soccer. Ward and Williams (2003) found that a battery of visual acuity, depth sensitivity, and peripheral awareness measures were unrelated to skill in a sample of elite and sub-elite soccer players. Starkes (1987) found that elite female field hockey players had a significantly slower simple reaction time than sub-elite and novice players. The mixed results in the academic literature outlines the importance of more closely examining the assumed correlations between athleticism and future performance before they inform decisions about player value. While it may be that elite basketball players perform better on basic athletic ability tasks, the theoretically important question is whether these athletic attributes uniquely predict performance at the next level across a large sample of college players.

Past research on the NBA draft

Predicting NBA performance and NBA draft behavior has not been of much interest to the scholarly community until recent years. An early study failed to find any objective performance based predictors of draft order (Kahn & Sherer, 1988). Coates and Oguntimein (2010) showed that college performance predicted NBA careers and draft order. In contrast to Kahn and Sherer (1988), this study showed that performance variables including points, rebounds, and assists predicted draft order. They also showed that draft order was related to NBA opportunity as well as performance; and that a player's college statistics predicted their NBA equivalents. Finally, Coates and Oguntimein (2010) showed that the quality of the college conference changed how players were assessed. Research has also shown that better draft status, independent of performance, provides more sustained opportunities for players (Staw & Hoang, 1995).

A recent study by Berri, Brook, and Fenn (2011) showed that the most important performance variable for predicting draft order was scoring. There are also several other performance variables used to predict draft order like assists, blocks, and shooting percentage; as well as non-performance factors like height, age, and team success that particular season. Multiple measures predicted NBA performance including points (negatively), rebounds, steals, shooting efficiency, and team success (negatively). In their equation, height did not predict performance and age was only related for players in their fifth year. The youth advantage, which suggests that age predicts relative success, has also been confirmed by separate analyses (Rodenberg & Kim, 2011).

These studies allow us to draw several conclusions. First, draft status is correlated with opportunity and it is likely that its predictive validity is exaggerated. Second, predictors of draft status and NBA performance overlap in many ways. For example, college performance measures and age predict both draft status and performance, but differ in that other variables such as height and team success predict draft status but not NBA performance. With the exception of Coates and Oguntimein (2010), most studies use many related variables to predict a single measure of NBA performance. This introduces collinearity that may be obscuring some theoretically and practically important relations.

The current study

In the current study we attempt to identify the quantifiable traits that define untapped potential as players advance in their basketball careers. The NBA draft offers a strong pool of data to examine. Before every draft, the NBA holds a combine where various body size and athletic measurements are taken. The NBA selects the players who are invited and typically the vast majority of people who are drafted from college attend. Since the NBA takes the time to provide these measures to their teams we can infer that they are believed to be helpful in spotting untapped potential. Websites that specialize in updating interested fans now publish combine data publicly every year. For most players there is also college performance data available, which is measured in similar metrics as those used in the NBA. Additionally, the data allows for a more complete comparison of non-basketball performance measures than in previous studies. Of the previous studies reviewed, none had athleticism measures or variables like arm span, which might be more useful than height alone. It should be noted that this study necessarily uses an elite sample. For instance, in the first year of our sample only about 6% of eligible division I basketball players attended the combine. However, the sample is drawn from the population of interest, which includes those who are skilled enough to have the opportunity to be considered by the NBA as a draft prospect.

This study will construct growth mixture models of the first 3 years of NBA success. A growth mixture model classifies people into latent classes based on information about where they start and how they change (i.e. intercept and slope) (See Nylund, Asparouhov, & Muthen, 2007). We predict that there are multiple classes of players who enter the NBA draft; players with significantly different initial performance and/or growth. We will examine the first three years of NBA performance, a critical period after which an NBA team must decide whether or not to offer a contract extension. We will attempt to predict class membership with a series of performance, demographic, anthropometric, and athletic variables. We hypothesize that past performance, age, and the university's history of producing NBA players (college quality) will predict class membership, but anthropometry and athleticism will not. It is important to clarify that we do not argue that anthropometry and athleticism are unimportant to NBA performance, but instead, we propose that these variables have already expressed themselves through previous performance and thus do not uniquely predict NBA player class membership. In a recent article Detterman (2014) noted a lack of correlation within the NBA between height and performance but the large mean height difference between NBA players and the general population. Detterman argued that the mean difference is evidence for the importance of height even without a correlation within restricted elite sample. According to this perspective, as the sample is broadened, the correlation between performance and height should increase. Our sample, which will include players who never made the NBA, and those who dropped out of the league quickly, should have a much better chance of discovering such a relation than analyses with only super-elite players. Even the study by Berri, Brook, and Feen (2011) which found height to be a non-significant predictor was more restricted than our study as it only used players that were drafted and did not include players that were considered for drafting, who often have successful basketball career in the NBA or internationally.

Method

Collection of data

Three sets of data were collected and combined, size and athletic measures of each player were collected from an NBA combine performance database at http://www.draftexpress.com. College performance variables were collected from http://www.sports-reference.com, and NBA performance and position data were collected from http://www.basketball-reference.com.

Variables used

There are many anthropometric and physiological measures taken at the combine. However, we constrained the number used in our analyses for several reasons. First, the variable must have been consistently measured. A few variables such as hand size and body fat were not available for most players. Additionally, to reduce collinearity issues, we selected one variable among those that are measured multiple ways. For example, no step vertical leap and maximum vertical leap, agility test and sprint test, and finally height and arm span. From these, we selected no step vertical leap, agility and arm span. Since we hypothesized that these variables would not predict NBA performance, we used the variable that was most correlated with players' summed NBA performance over their first three years so that the variable selection of rejecting the null and failing to support our hypothesis.

For our measure of performance we used win shares, a reliable metric of all-around performance that correlates well with other aggregate metrics of performance (Winston, 2009). The win shares metric uses a mixture of statistical efficiency measures, including how often players are involved in plays and a team's defensive efficiency with the player on the court (Winston, 2009). Win shares is not perfect because the defensive measure is influenced by the quality of a player's teammates. However, it is reliable. In our sample, year-to-year correlations ranged from .77 to .85. It also has strong face validity; the top five players in the 2012 season: Lebron James, Kevin Durant, Chris Paul, James Harden, and Russell Westbrook, were all-stars. Moreover, the players ranked in the top 25 in NBA history in win shares have all been selected to the Basketball Hall of Fame. Another advantage of using win shares is that the statistic is available for college players. Those players who attend the NBA combine and did not play every (or any) year in the NBA were given a zero value, as were players who scored lower than zero.

In sum, predictor variables used in the final analysis were: player position (guard, forward, center), age at the beginning of the next NBA season, the player's college win shares, college quality, NBA combine agility, combine no-step vertical leap, combine arm span, and combine weight. Position was coded 0 for guard, 1 for forwards, and 2 for centers. We would have preferred to code 1–5 (point guard, shooting guard, power-forward, small forward, and center), but the necessary data was not included for all players. The agility test consists of running laterally between cones; faster times are better. No-step vertical leap is equivalent to a standing vertical leap. The player is required to jump as high as possible without gaining any horizontal momentum. Arm span is a measure of length of the distance from a person's left to right fingers when their arms are extended, fully outward and parallel to the ground.[1]

Plan for data analysis

Data analysis proceeded in four stages. For the first step, four growth mixture models were fit. These models allowed only one, two, three, or four classes. In each model the class membership was predicted by the variables listed above. AIC and BIC was used for model selection using the smallest-is-best criterion. In the second stage we used the class membership probabilities and the class values to project each player's three-year career and correlated our projection with the actual data.

[1] It is fair to question how reliable the measures we extracted are. There are a couple of ways we checked this. First multiple sources, including the NBA, publish draft combine data. We found that, excepting rounding differences, those values obtained from draft express correlated perfectly with those published by the NBA. Secondly several of the measures have a direct analogue, as pointed out earlier. For instance the correlation of height with armspan was .84 and the correlation of no step vertical leap with their maximum vertical was also .84. While agility is a unique task, it did correlate .50 with sprinting speed, .43 with vertical leap, and .48 with height as would be expected. These are all obviously different measures so they can be taken as lower bounds for the reliability of the measures. Finally we had 26 people who attended the combine twice. In the analysis we only used the final combine results. We can calculate a lower bound on test retest reliability. These estimates are obviously conservative as these are young men who may not have fully developed physically or athletically and in one year compared to the next. The correlations were as follows all are significant $p < .01$: for armspan $r(24) = .85$, for agility $r(17) = .65$, and for no step vertical leap $r(17) = .71$. Together the consistency across sources and the logical correlation pattern within the data suggest we harvested meaningful and accurate data from our source.

To clarify, let us consider an example. Assume there are two classes, the first with an intercept of 2 win shares and a slope of .5 win shares and a second with an intercept of 1 win shares and a slope of 0 win shares. Next, assume the class membership probability (a measure of how close to various classes the subjects data is) was .75 for the first class and was .25 for the second class. For the first year our predicted value would be (2 win shares*.75) + (1 win shares*.25) = 1.75 win shares for the first season. Each year after that would increase by the following value: (.5 win shares*.75) + (0 win shares *.25) = .375 win shares increase each subsequent season. So for the second year we would predict a value of 2.13 win shares and 2.50 win shares for the third year. Correlations will be reported for year 1, 2, and 3. This will be done with the data from the 2001–2006 draft classes. We will then force a forecast model using the exact parameters of the 2001–2006 data to fit the data for the 2007–2009 drafts (the last drafts with three subsequent years available for analysis before the strike shortened year of 2012) we will predict NBA performance with this forecast model in exact same manner. The growth mixture model will be conducted using the MPlus statistical analysis software (Muthén & Muthén, 1998–2011).

Results

Descriptive statistics and missing data for 2001–2006 draft classes

Of the 408 players available who played college basketball in the United States and attended the combine from 2001 to 2006, only 32 had incomplete data. The most common omission was agility, which 29 players did not complete. Players who had complete data did not differ significantly from players with incomplete data on any of the predictor variables, although the players with complete data had slightly more win shares in the first three years of their NBA career. Given that the players with incomplete data were less skilled, we will analyze only the complete cases particularly because the pivotal athletic variables did not have acceptable data coverage in athletic variables for the missing data group.

Our sample contributed an average of 0.63 win shares their first year ($SD = 1.38$), 0.96 their second year ($SD = 1.99$), and 1.22 their third year ($SD = 2.57$). In their last year of college they had contributed an average of 4.57 win shares ($SD = 1.81$). The average university win shares, which consist of a sum of the win shares produced by draftees from a particular school over the previous decade, was 52.01 ($SD = 63.86$). The average age of our sample when entering the NBA was 22.72 ($SD = 1.31$). The average weight was 218.01 pounds ($SD = 27.22$). While height is not used as a predictor, the average height is 77.53 inches ($SD = 3.45$) or 6′5″, which is clearly taller than the national male average of approximately 5′9″. The average arm span is 82.31 inches ($SD = 4.05$). The average time on the agility drill was 11.48 s ($SD = .61$) and the average no-step vertical leap was 28.92 inches ($SD = 3.04$). Two hundred and eight people in our sample were not drafted.

Growth mixture analysis

An analysis of the model fit revealed that the one-class model had an AIC of 21847.71 and a BIC of 22056.29. The two-class model fit the data significantly better with an AIC of 3725.92 and a BIC of 3800.58. The three-class model was clearly superior to the two-class model with an AIC of 3446.16 and a BIC of 3564.05. The four-class model did not improve on the three-class model having an AIC of 3714.40 and a BIC of 3875.51 demonstrating that the three-class model best fit the available data.

The analysis identified 308 of the 376 players as most likely to belong in the first class, 18 as being most likely to belong in the second class, and 50 as being most likely to belong in the third class. Class one had an intercept of 0.17 win shares ($p = .001$) and a slope of 0.15 win shares ($p < .001$). Class two had an intercept of 0.69 win shares ($p = .01$) and a slope of 0.08 win shares ($p = .31$). Class three had an intercept of 3.6 win shares ($p < .001$) and a slope of 1.3 win shares ($p = .001$). Likelihood of belonging to the second class instead of the first class was predicted positively by position and no-step vertical, and negatively by age, college win shares, quality of college basketball program, agility (which would mean more agile players were more likely to be members of class two), arm span, and weight ($p < .002$ for all predictors). The only variables that significantly differentiated class three from class one were age, college skill, and college quality (see Table 1 for complete data). In sum, the variables that are most important for predicting NBA success are age, college win shares, and past college success.

Predicting NBA success

Our model correlated with NBA performance significantly $r(374) = .78$, $p < .001$ for the first, $r(374) = .84$, $p < .001$ for the second, and $r(374) = .71$, $p < .001$ for the third year. Fig. 2 displays a scatterplot showing relationship between our model prediction and observed NBA performance. For a comparison, we analyzed the spearman's correlations between draft order and NBA win-shares and assigned all undrafted players the same value lower than any drafted player. The Spearman correlations were $rs(374) = -64$, $rs(374) = -.70$ for the second year, and $rs(374) = -.67$ for the third year (all years $p < .01$). In all cases the Spearman coefficient was higher than the Pearson correlation.

To show that our latent class model not only fit the data better but also allowed us to predict the criterion better we did conduct a multiple regression with our predictor variables regressed on NBA performance for each year. For year 1 the multiple correlation was $r(367) = .34$, $p < .001$, for year 2 the multiple correlation was $r(367) = .35$, $p < .001$, and for year 3 the multiple correlation was $r(367) = .40$, $p < .001$.

Forecasting model

Since all parameters were fixed we cannot report any significances. AIC was 2330.47 and BIC was 2330.47. Overall there were 159 people placed in class one, 12 in class two, and 34 in class three. The important question is how well does the established model

Table 1
Summary of changes in class probabilities due to different variables compared to class 1 ($N = 376$).

Variable	*B*	SE	C.R.	*p*
Class 2				
Age	−44.21	.23	−190.66	<.01
College win shares	−35.23	.30	−116.15	<.01
College quality	7.76	<.01	1955.62	<.01
Position	15.90	.68	23.46	<.01
Agility	−75.17	.60	−125.27	<.01
No step vertical	9.17	.09	106.04	<.01
Weight	−.73	.02	−437.48	<.01
Arm span	−.36	.11	−3.39	<.01
Class 3				
Age	−1.02	.23	−4.41	<.01
College win shares	1.11	.30	3.63	<.01
College quality	.02	<.01	3.95	<.01
Position	−.33	.68	−.49	.63
Agility	−.53	.60	−.89	.37
No step vertical	.04	.09	.49	.62
Weight	−.01	.02	−.35	.73
Arm spsan	.02	.11	.17	.87

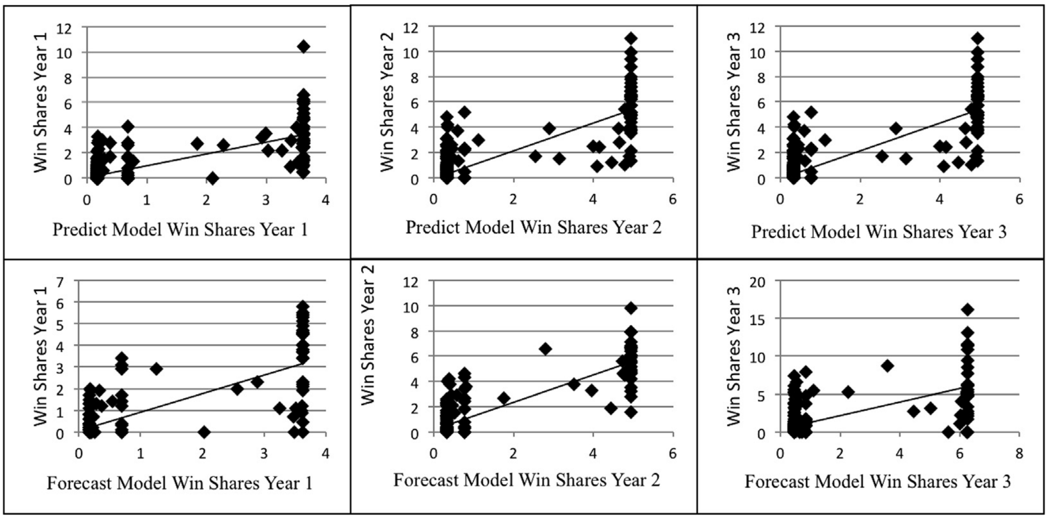

Fig. 2. Scatter plots of models' predicted win shares against NBA win shares. Top is the prediction model and bottom is the forecast model.

predict NBA success. The correlations of our forecast with subsequent NBA success are as follows: $r(203) = .78$, $p < .001$ for the first season, $r(203) = .84$, $p < 001$ for the second season, and $r(203) = .64$, $p < .001$ for the third season. We, again, compared our forecast with draft order given all undrafted players the same value lower then any drafted player. The Spearman correlations were $rs(348) = -.52$, $rs(374) = -.66$ for the second year, and $rs(374) = -.68$ for the third year (all years $p < .01$). Reference Fig. 1 to contrast our model prediction to the observed NBA performance. To compare, the multiple correlation of our predictor variables with each year was as follows $r(198)-.41$, $p < .001$, $r(198) = .54$, $p < .001$ for year 2, and $r(198) = .57$, $p < .001$ for year 3.

Predicting NBA draft status

We also conducted a hierarchical multiple regression predicting NBA draft order with our model variables. For this analysis, we combined the datasets after confirming that the confidence intervals overlapped for all effects (at least 40% overlap in all cases). Table 2 gives the full model prediction and effect sizes. The variables we found to be predictive of performance, player age, players' college win shares, and college quality, were entered in step 1 and were significantly related to draft status; $F(3, 578) = 150.96$, $p < .001$. The athletic and anthropometric variables were also significant $F(5, 573) = 3.31$, $p = .01$. The full model predicting NBA draft status shows that draft order is largely related to college win shares (negatively) and age (positively). By convention, both of these effect sizes would be considered medium (Cohen, 1988). College quality also predicted draft order, although only weakly so.

The variables predictive of draft status are of theoretical interest because our model showed that the three variables were equally discriminative of class as measured by the critical ratio (C.R. = beta/standard error), which suggests that college quality is not being completely taken into account by NBA teams. Additionally, we found that better agility and longer arms was associated with improved draft status. When removing step one in the regression model, the removal was significant, $F(1, 573) = 135.00$, $p < .001$, as were the remaining athletic and size variables $F(5, 576) = 8.62$, $p < .001$. The only interpretive difference in this case was that playing in a more inside position now predicted worse draft status $t(1, 576) = 3.02$, $p = .001$, $f^2 = .02$.

Discussion

The current study makes several important contributions to our understanding of talent identification. First, contrary to the widely held views of untapped potential, we did not find a strong role for physical attributes and athleticism in distinguishing productive NBA players from unproductive ones after controlling for previous performance. This finding is different than previous studies because we forecasted future success with a wider sample and gave a strong hearing to these measures by avoiding collinearity while choosing the variables most likely to predict NBA performance in a full model. These physical attributes only differentiated players who would make functionally no impact in the NBA from those

Table 2
Summary of hierarchical multiple regression predicting NBA draft status. All estimates are shown for step 2. Step 3 removes the variables entered in step 1. Negative relationships denote better draft status ($N = 582$).

Variable	B	SEM	β	f^2	R^2	ΔR^2
Step 1: Enter Age, College win shares, and College quality					.44**	
Step 2: Enter Anthropometric and Athletic variables					.46**	.02**
Age	10.61	.78	.44**	.32		
College win shares	−7.10	.61	−.37**	.23		
College quality	−.05	.02	−.10**	.02		
Position	3.99	2.40	.08	<.01		
Agility	4.90	2.18	.09*	.01		
No step vertical	−.21	.37	−.02	<.01		
Weight	.10	.06	.01	<.01		
Wing span	−1.39	.42	−.17**	.02		
Step 3: Remove variables in step 1					.07**	−.39**

Note * = $p < .05$ and ** = $p < .01$.

who would continue in the league as reserve players. The only variables that predicted NBA success were age, players' college win shares, and college quality. Past studies have looked at height, but our study is the first that we are aware of that also examines athletic measures, which are also thought to predict untapped potential. Our analysis extends upon past analyses by identifying significant predictors of draft status and relating those predictors directly to subsequent performance in the NBA. Additionally we showed that a model based on prediction of latent class membership could predict and forecast future NBA performance better than draft status or a simple multiple regression model. This suggests latent class analysis may be a useful tool for predicting professional sport performance. Finally, we showed that NBA draft order appears to value anthropometric and athletic variables despite those variables not being reliably related to NBA success. Thus, this series of analysis largely conformed to our theoretical expectations. As predicted by current expert performance theories, skill transitions through a series of stable states and predictable developmental patterns, given that deliberate practice is occurring (Ericsson et al., 1993). Within this theoretical perspective, college performance should strongly predict NBA performance and identifying factors that predict performance above and beyond previous performance is problematic.

It is also interesting that playing an interior position is negatively related to draft status when step one is removed, because it suggests that centers come into the league with a weaker background as measured by age, college win shares, and college quality, and that this is appropriately accounted for by NBA general managers. This goes against the lay wisdom that NBA general managers draft centers earlier than they should. The results also show that general managers factor arm span and agility into their draft day decisions.

Additionally, younger players whose skill is comparable to older players at draft time may have accumulated a similar amount of deliberate practice (Ericsson et al., 1993) to reach their current level of skill, and thus they are better situated to continue improving. The predictive power of college quality can also be explained in a similar way. Those players with goals of playing in the NBA are more likely to organize their lives around receiving the best instruction (Bloom, 1985) these effects also likely have benefited the younger players. In current models of expertise, the skill of the coach is thought to be very important because quality coaches organize appropriate activities and give proper feedback, which are both essential components of deliberate practice (Ericsson et al., 1993). This view contrasts with simpler models that operationalize training effects as only a factor of the quantity of practice (see Hambrick, et al., 2013).

Limitations

The first limitation of this study is that not all data was collected for all players. In fact, in recent years increasing numbers of players appear to be skipping the athletic tests in the combine. Given that our model of that data was a forecast model, it did not affect our parameter estimates. However, variables such as hand size and body fat, may have been predictive of class, but were simply too sparse to analyze. Additionally, each team holds tryouts to individually scout players. We cannot say how well the data collected at these types of events predicts success or if any of those collected variables might identify untapped potential. We were not able to project success for two important populations: international players and players entering the NBA from high school, so our results do not generalize to these groups. Also, our measure of skill is imperfect, if we created more detailed models based on more diverse components of basketball such as passing, defense, rebounding, shooting, and creating a shot for others, we may have found different results. Finally, given the correlational nature of our design, many outside factors may also predict players' NBA performance. These could include coaching, team-level organization and management, random injuries and team social dynamics. These factors may be noise to the year over year effects. More detailed data is becoming available all the time in basketball, which may allow these types of hypotheses to be effectively tested in the future.

The role of experience in NBA performance

As mentioned earlier, a common feature of those who go on to become experts is the early recognition of perceived potential by a parent or coach (Bloom, 1985). While, systematic reviews show that early identification of talent is often confounded with variables such as relative age (Musch & Grondin, 2001). These early advantages can lead to permanent and compounding advantages (Merton, 1968). For instance, in chess and music, the best performers in a cohort will typically be those that started earliest (Ericsson et al., 1993; Gobet & Campitelli, 2007). The mechanism of this advantage has not been identified in chess, but in music there is evidence that early training leads to neurological changes (Bengtsson et al., 2005). Similarly, in sports, early training can lead to advantageous anatomical adaptations that cannot be acquired after certain ages (Ericsson et al., 2007).

Based on these findings, we would expect that younger players who are drafted would be from a population that started earlier and were identified as talented more often by coaches and caregivers. Players would have more (and likely higher quality) domain-specific practice than their peers and would therefore be more skilled (Ericsson et al., 1993). These individuals subsequently receive more effective coaching, better training opportunities in the off-season, earlier draft status, and more opportunities at all phases of development. They would also be the most likely to have the made the requisite physiological adaptations during possible critical developmental periods. Table 3 displays all of the correlations for measures used in this study for both the prediction and the forecast dataset. Table 3 shows that younger players are, on average, not better in college than older players. This combined with the fact that they also would have played fewer years in the college means that the college programs get less value from them. Due to high school scouting, coaches know which players are likely to leave after one or two years of play. However, as Table 3 shows, these players went to schools with a better history of producing NBA players and likely receive better coaching there.

To summarize, NBA scouts are looking for variables that predict untapped potential, which we have defined as the second component of talent, which is critical to identify prior to making a draft pick. We analyzed these proposed variables in a way designed to increase the likelihood that we could identify significant contributions without confounding our measures by using the variables collected by the NBA for that purpose and being aware of collinearity. These physical, or "untapped potential" variables did not predict membership in the only class with successful NBA careers. Instead, NBA productivity was predicted by previous performance, as well as the quality of the college attended, which likely represents better training. Consistent with our hypotheses, players' age predicts performance in a way that suggests Matthew Effects. There also appears to be some compensation for height in our data; players with longer arms do not jump as high nor are they as agile as those with shorter arms (see Table 3). This suggests that physical thresholds may exist, specifically for how much court can be covered by a player and for their ability to play above the basket,

Table 3
Correlation Matrix of variables used in this study for 2001–2006 drafts below the axis ($n = 376$) 2007–2010 above the axis ($n = 206$).

		1	2	3	4	5	6	7	8	9	10	11	12	13
1.	Win shares year 1	1.00	.67*	62*	.80*	.33*	−.10	−.22*	.07	−.39*	−.04	.15*	−.02	−.02
2.	Win shares year 2	.77*	1.00	.77*	.92*	.42*	−.12	−.33*	.06	−.51*	−.09	.19*	−.06	−.06
3.	Win shares year 3	.82*	.89*	1.00	.94*	.45*	−.05	−.36*	.13	−.51*	−.04	.18*	.01	.02
4.	Win shares year 1−3	.90*	.94*	.97*	1.00	.46*	−.09	−.36*	.10	−.54*	−.06	.19*	−.02	−.02
5.	Last college season win shares	.39*	.40*	.39*	.42*	1.00	.08	−.14	.01	−.43*	−.10	.10	.14*	.13
6.	Position	.01	−.02	−.03	−.02	.03	.03	−.03	−.08	.05	.44*	−.29*	.72*	.70*
7.	Age	−.30*	−.27*	−.29*	−.30*	−.13	.20*	1.00	−.15*	.63*	.02	−.07	−.15*	−.09
8.	NBA win shares produced by college	.26*	.19*	.17*	.21*	.07	.00	−.17*	1.00	−.16*	−.13	.02	−.03	−.03
9.	Number drafted	−.56*	−.57*	−.56*	−.60*	−.45*	.04	.44*	−.23*	1.00	.15*	−.18*	−.12	−.02
10.	Agility	−.09	−.14*	−.08	−.11*	−.12*	.05	.08	−.09	.13*	1.00	−.30*	.47*	.44*
11.	No step vertical leap	.11*	.11*	.11*	.12*	.02	−.27*	−.15*	.02	−.09	−.34*	1.00	−.15*	−.33*
12.	Arm span	.07	.04	.05	.06	−.14*	.69*	.09	.05	−.09	.46*	−.14*	1.00	.70*
13.	Weight	.02	.01	−.00	.01	.06	.71*	.15*	.06	−.01	.45*	−.24*	.73*	1.00

Note. Correlations greater than or equal to $\pm .19$ are significant at $p < .05$ and correlations $\pm .24$ are significant at $p < .01$. *= $p < .05$.

both of which can be accomplished by being tall or by being exceptionally agile.

Matthew effects such as the ones observed here are troublesome for interpretation both for our model and for arguments such as Detterman's (2014). If, due to stereotypes or lay psychology, a variable is overvalued by the people deciding who gets an opportunity at each level it will obviously raise the mean level of that variable past what would be optimal. This suggests that the mean differences between NBA players and the general population, noted both by this study and Detterman (2014) might be at least partially the result of factors unrelated to objective superior skill. At the same time if tall but inferior players were given unearned opportunities it would also reduce the correlation between size and performance. If basketball (at all levels of competition) functioned as a perfect meritocracy it is possible that the average height would be shorter and that there would be a correlation between height and performance. The only way to answer these types of questions is to collect strong data early in the development processes to model performance at younger ages. Future studies should use our method of projecting both opportunity and success at the next level to better create a model of success in basketball while differentiating between earned and unearned opportunities.

The data fits the predicted pattern based on our understanding of environmental influences, but does not exclude a accounts of individual differences in talent. For example, age could be a proxy for a latent talent variable. This is a general problem with talent explanations because testing what are believed to be the most likely manifestations of talent does not rule out all others. In a domain like the NBA, we cannot ignore that selection effects have already occurred based off likely biomechanical advantages (e.g. NBA players are much taller and have much longer arms than could be possible from a randomly selected sample). To clarify, our argument is that these physical characteristics have already contributed to achievement and training opportunities by the time the player reaches the NBA and thus does not differentiate player class at the professional level. The variables that continue to predict NBA performance are proxies for current performance and quality of training environments, which follows well-established patterns from many domains of expertise.

References

Associated Press. (2012, June 29). *Hornets draft Anthony Davis, Austin Rivers. Sports illustrated*. Retrieved from http://sportsillustrated.cnn.com/2012/basketball/nba/wires/06/28/2030.ap.bkn.draft.hornets.5th.ld.writethru.1176/index.html.

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., & Ullén, F. (2005). Extensive piano practicing has regional specific effects on white matter development. *Nature Neuroscience, 8*, 1148–1150.

Berri, D. J., Brook, S. L., & Feen, A. J. (2011). From college to the pros: predicting NBA amateur player draft. *Journal of Productivity Analysis, 35*, 25–35.

Berri, D. J., & Schmidt, M. B. (2006). On the road with the National Basketball Association's.superstar externality. *Journal of Sports Economics, 7*, 347–358.

Bloom, B. S. (1985). Generalizations about talent development. In B. S. Bloom (Ed.), *Developing talent in young people* (pp. 507–549). New York: Ballantine Books.

Charness, N., Tuffiash, M., Krampe, R., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology, 19*(2), 151–165.

Coates, D., & Oguntimein, B. (2010). The length and success of NBA careers: does college production predict professional outcomes? *International Journal of Sports Finance, 5*, 4–26.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Detterman, D. K. (2014). Introduction to the intelligence special issue on the development of expertise: Is ability necessary? *Intelligence*. http://dx.doi.org/10.1016/j.intell.2014.02.004. Advance online publication.

Elo, A. E. (1978). *The rating of chessplayers, past and present*. London: Batsford.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406.

Ericsson, K. A., Roring, R. W., & Nandagopal, K. (2007). Giftedness and evidence for reproducibly superior performance: an account based on the expert performance framework. *High Ability Studies, 18*, 3–56.

Finnan, B. (2013, July, 13). *Cavaliers: Mike Brown concocting plans for top pick Anthony Bennett*. The News-Herald. Retrieved from http://www.news-herald.com/general-news/20130716/cavaliers-mike-brown-concocting-plans-for-top-pick-anthony-bennett.

Fryar, C. D., Gu, Q., & Ogden, C. L. (2012). *Anthropometric reference data for children and adults: United States, 2007–2010*. In *Vital Health Stat* (Vol. 11) (p. 252). National Center for Health Statistics.

Galton, F., Sir (1869/1979). *Hereditary genius: An inquiry into its laws and consequences* (Originally published in 1869). London: Julian Friedman Publishers.

Gobet, F., & Campitelli, F. (2007). The role of domain-specific practice, handedness, and starting age in chess. *Developmental Psychology, 42*, 159–172.

Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2013). Deliberate practice: is that all it takes to become an expert? *Intelligence*. http://dx.doi.org/10.1016/a0029082/j.intell.2013.04.001. Advance online publication.

Hausman, J. A., & Leonard, G. K. (1997). Superstars in the National Basketball Association: economic value and policy. *Journal of Labor Economics, 15*, 586–624.

Kahn, L. M., & Sherer, P. D. (1988). Racial differences in professional basketball players' compensation. *Journal of Labor Economics, 6*, 40–61.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*, 56–63.

Moxley, J. H., & Charness, N. (2013). Meta-analysis of age and skill effects on recalling chess positions and selecting the best move. *Psychonomic Bulletin & Review*. http://dx.doi.org/10.3758/s13423-013-0420-5. Advance online publication.

Musch, J., & Grondin, S. (2001). Unequal competition as an impediment to personal development: a review of the relative age effect in sport. *Developmental Review, 21*(2), 147–167.

Muthén, L. K., & Muthén. (1998-2011). *Mplus user's guide* (6th ed.). Los Angeles: Muthén & Muthén.

Nylund, K. L., Asparouhov, T., & Muthen, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a monte carlo simulation study. *Structural Equation Modeling, 14*, 535–569.

Rodenberg, R. M., & Kim, J. W. (2011). Precocity and labor market outcomes. Evidence from professional basketball. *Economics Bulletin, 31*, 2185–2190.

Starkes, J. L. (1987). Skill in field hockey: the nature of the cognitive advantage. *Journal of Sport Psychology, 9*(2), 146–160.

Staw, B. M., & Hoang, H. (1995). Sunk coasts in the NBA: why draft order affects playing time and survival in professional basketball. *Administrative Science Quarterly, 40*, 474–494.
Walters, C., & Williams, T. (2012). To tank or not to tank? Evidence from the NBA. In *MIT Sloan Sports Conference March. Boston, MA*.
Ward, P., & Williams, A. M. (2003). Perceptual and cognitive skill development in soccer: the multidimensional nature of expert performance. *Journal of Sport & Exercise Psychology, 25*(1), 93–111.
Winston, W. L. (2009). *Mathletics*. Princeton, NJ: Princeton University Press.

Web references

NBA & ABA Player Directory. (2013, September). Sports Reference LLC. Retrieved from http://www.basketball-reference.com/players/.
NBA Pre-Draft Measurements. (2013, September). Draft Express. Retrieved from http://www.draftexpress.com/nba-pre-draft-measurements/.
Player Index. (2013, September). Sports Reference LLC. Retrieved from http://www.sports-reference.com/cbb/players/.