



Eye on the ball: the relationship between sensorimotor abilities and on-field performance in professional baseball

Paper Track: Baseball
Paper ID: 5425

Abstract

Baseball players must be able to see and react in an instant, yet it is hotly debated whether superior on-field performance is associated with superior sensorimotor abilities. In this study, we compare sensorimotor abilities, measured through eight psychomotor tasks comprising the Nike Sensory Station battery, and game statistics in a sample of 252 professional baseball players. For this purpose, we develop a series of Bayesian hierarchical latent variable models that enable us to compare statistics across multiple professional baseball leagues. Within this framework, we find that sensorimotor abilities are statistically significant predictors of on-base percentage, walk rate, and strikeout rate, accounting for confounding variables such as age, position, and league. We find no such relationship for either slugging percentage or fielder-independent pitching. The pattern of results suggests performance contributions from both visual-sensory and visual-motor abilities and indicates that sensorimotor screenings may be useful for player scouting.

1. Introduction

Ted Williams, one of the most legendary baseball players of all time, once said, “I think without question the hardest single thing to do in sport is to hit a baseball.” Advances in sport science continue to validate Williams’ claim; hitting a pitched baseball places incredible demands on athletes’ visual systems. We now know that Major League Baseball (MLB) pitches move at speeds near the processing limits of vestibular-ocular tracking [1], leaving the batter with mere milliseconds to decipher the pitch, project its trajectory, decide to swing, and coordinate the timing and trajectory of a 2.5-inch diameter bat. The immense difficulty of this task is underscored by the fact that players who hit successfully on less than a third of their at-bats can receive nine-figure contracts in today’s free-agent market.

Pitching, while equally demanding, draws upon a fundamentally different skill set. Pitchers attempt to deny batters effective contact with the ball while projecting it through the strike zone sixty feet away. Despite the need to visualize the strike zone, it has been argued that motor demands, such as controlling the speed, spin, and location of the ball, are more important for pitching success than visual requirements [2].

Given the substantial role of visual and motor demands in baseball (henceforth called “sensorimotor skills”), there has been a concerted effort to determine which elements of the perception-action cycle contribute to successful baseball performance [3]. However, the combination of noisy game statistics and costly sample acquisition makes inferring meaningful relationships difficult.



Between 2011 and 2015, the Nike Sensory Stations were developed and utilized as a tool to quantitatively evaluate athlete visual and motor skills. Participants filled out a registry of information about themselves and completed a battery of nine visual-motor tasks administered under standardized conditions with video instructions and conducted by trained and certified administrators. Data from these assessments were maintained on a central database and used to provide sport-specific normative information to individuals about their relative abilities and to monitor learning when coupled with sensorimotor training interventions.

Past research with the Sensory Stations has demonstrated that the battery of tests is reliable [4, 5] and cross-validated [6], with some tasks demonstrating linear improvements with practice over multiple sessions [7]. Improved performance on this battery has been seen following sports vision training interventions [8] and has been linked to baseball batting expertise, with professional baseball hitters showing better performance on measures of visual sensitivity relative to pitchers [9]. Furthermore, reduced performance on these tasks has been associated with an increased likelihood of sustaining head impacts during practices and games for American collegiate football players [10]. In addition, Poltavski and Biberdorff [11] found that better performance on measures of dynamic visual acuity and visual motor control accounted for nearly 70% of the variability in goals scored over two seasons in a sample of 19 men's and 19 women's collegiate hockey players. Collectively, past research [reviewed by 12] suggests that this battery may serve as a useful tool for understanding human performance, warranting further investigation into the sensorimotor characteristics of athletes and their relation to performance outcomes.

2. Methods

2.1. Data

In the current study, Sensory Station assessments from 252 professional baseball players collected in 2012 and 2013 were compared to game statistics to evaluate the relationship between sensorimotor skills and baseball production. For each player, game statistics from the season after testing were acquired along with information about their league(s) of participation. All data were shared with the research team under a secondary-data protocol approved by the Duke University Institutional Review Board [IRB B0706]. Under this protocol, all data were collected for "real world use," without informed consent, and shared with the research team after removal of all protected health information (PHI). As such, these data conform to U.S. Department of Health and Human Services, "Regulatory considerations regarding classification of projects involving real world data," [13].

2.1.1. Sensorimotor Assessments

The Sensory Stations consist of a battery of nine computerized sensorimotor tasks, each designed to evaluate a specific facet of a participant's visual-motor abilities. Each task is briefly described below and is depicted visually in **Figure 1**. More detailed descriptions of the tasks, along with the behavioral performance distributions on these measures can be found in the **Appendix**.

- The **Visual Clarity** task measures visual acuity for fine details at a distance.
- The **Contrast Sensitivity** task measures the minimum resolvable difference in contrast at a distance.



- The **Depth Perception** task measures how quickly and accurately participants are able to detect differences in depth at a distance using liquid crystal glasses.
- The **Target Capture** task measures the speed at which participants can shift attention and recognize peripheral targets.
- The **Near-Far Quickness** task measures the number of near and far targets that can be correctly reported in 30 seconds.
- The **Perception Span** task measures the ability to remember and recreate visual patterns.
- The **Eye Hand Coordination** task measures the speed at which participants can make visually-guided hand responses to rapidly changing targets
- The **Go/No-Go** task measures the ability to execute and inhibit visually guided hand responses in the presence of “go” and “no-go” stimuli
- The **Reaction Time** task measures how quickly participants react and respond to a simple visual stimulus.

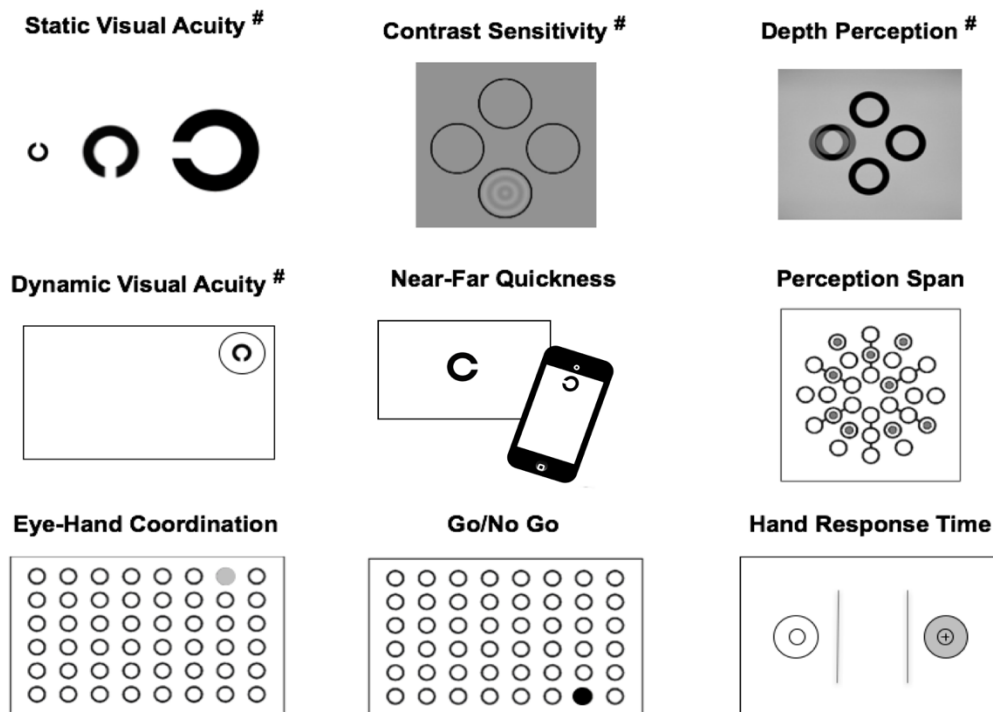


Figure 1: Illustrations of the nine perceptual and visuomotor tasks included in the Nike SPARQ Sensory Station battery. # indicates tasks performed under a staircase schedule.



2.1.2. Response Variables

The sensorimotor assessments performed by the Nike Sensory Stations serve as our best measurement of a player's underlying sensorimotor abilities. Similarly, a player's game statistics are strong indicators of his on-field performance. Although advanced metrics such as wOBA and wRC are excellent at quantifying overall offensive performance, we were not provided enough information under the secondary-data protocol to calculate these statistics. As such, in this study, we use on-base percentage (OBP), walk rate (BB%), strikeout rate (K%), and slugging-percentage (SLG) to measure the performance of batters. In addition, we use fielder-independent pitching (FIP) to measure the performance of pitchers. Below are brief descriptions of each of these statistics and our motivation for using them as response variables in our models.

On-Base Percentage measures a player's propensity to reach base. On-base percentage is defined as

$$OBP = \frac{Hits + Walks + Hit\ By\ Pitch}{At\ Bats + Walks + Hit\ By\ Pitch + Sacrifice\ Flies} \quad (1)$$

On-base percentage is a simple and widely used metric for player evaluation, since frequently reaching base gives the offense more opportunities to score runs. Players with high on-base percentages consistently make effective contact with the ball and draw walks. As such, on-base percentage offers a robust measure of player productivity and a useful statistic by which to evaluate the relationship between sensorimotor abilities and on-field performance.

Walk Rate measures a player's propensity to draw walks. Walk rate is defined as

$$BB\% = \frac{Walks}{Plate\ Appearances} \quad (2)$$

Players who routinely draw walks generally differentiate well between balls and strikes, forcing the pitcher to throw pitches that are easier to hit. Walk rate can also provide information about a hitter's underlying approach at the plate.

Strikeout Rate measures a player's propensity to strike out. Strikeout rate is defined as

$$K\% = \frac{Strikeouts}{Plate\ Appearances} \quad (3)$$

Strikeouts are an unequivocally negative outcome for the offense and should be avoided in an at-bat. Although some successful players have high strikeout rates, a high strikeout rate indicates that a batter struggles recognizing pitches or making contact with the ball. A player who strikes out frequently and walks rarely typically has a dim future in baseball.

Slugging Percentage measures a player's propensity to hit for power. Slugging percentage is defined as



$$SLG = \frac{Total\ Bases}{At\ Bats} \tag{4}$$

Slugging percentage makes use of the fact that not all hits are equally valuable. Although it is an imperfect metric (e.g. doubles are not worth twice as much as singles), it does a decent job of quantifying batting power. Sensorimotor abilities may have different effects on a batter’s ability to hit for contact and ability to hit for power.

Fielder-Independent Pitching measures a pitcher’s run prevention, independent of the ability of the defense behind him. FIP is defined in terms of only variables that cannot be affected by the ability of the defense behind the pitcher.

$$FIP = \frac{13 * HR + 3 * (Walks + HBP) - 2 * Strikeouts}{Innings\ Pitched} + FIP\ Constant \tag{5}$$

According to FanGraphs [14], “*Fielder Independent Pitching (FIP) measures what a player’s ERA would look like over a given period of time if the pitcher were to have experienced league average results on balls in play and league average timing.*” It is generally more stable than ERA, since it is a measurement that cancels out the effects of defense and luck. Sensorimotor abilities may be related to pitcher performance, and FIP represents one of the best metrics for quantifying pitcher performance in a game setting.

2.1.3. Sample Characteristics

Although data was obtained for 308 professional baseball players (149 batters, 159 pitchers), we only examine data for the players with more than 30 at-bats or more than 30 innings pitched to mitigate the statistical noise associated with low sample sizes. This yields a final analyzed data set of 252 players (141 batters, 111 pitchers). **Table 1** reports the distribution of age and positional category in this sample. Most of the players in the sample are young prospects between 20-25 years old, though there are older players in the sample who disproportionately play in the Major Leagues.

	Batters	Pitchers
Age		
Mean (SD)	22.7 (3.9)	23.7 (3.6)
Min - Max	17 - 37	18 - 39
Position		
# Catchers	19	
# Infielders	65	
# Outfielders	57	

Table 1: Sample characteristics



Not all professional leagues are equal. The level of competition in Major League baseball significantly outclasses that of AA baseball, for example. Players in our sample played in leagues ranging from Rookie League to Major League Baseball, which makes player comparison more challenging. **Table 2** displays the number of players in our sample who play in each league. Note that some players play in more than one league.

	Rookie	A	Adv. A	AA	AAA	Majors
Batters	63	18	33	17	23	14
Pitchers	29	17	24	21	17	13

Table 2: Distribution of Leagues by Player Type

2.2. Statistical Models

We fit separate models for the five response variables. The models use a common set of predictors. For any model, all parameters are estimated using only the data for that model. For convenience, we use a common notation across models when describing the models, so α_j in the OBP model will be distinct from α_j in the SLG model, for example.

Binomial Response

Since OBP, BB%, and K% are defined as the number of successes divided by the number of opportunities, we use a binomial response for these three variables. Without loss of generality, we present the model for OBP. Let OB_{ij} denote the number of times that player i reached base in league j out of N_{ij} opportunities between 2012 and 2013. We treat each OB_{ij} as a realization of a random sample, with the player's true on base percentage equal to p_{ij} . Each p_{ij} is a function of the degree of difficulty of getting on base in league j , as well as the player's latent, on-base ability parameter A_i . Each A_i is a function of variables X_i that include the player's sensorimotor abilities, a set of indicator variables for position, and age. Putting it together, we have the Bayesian multilevel model [15]

$$OB_{ij} \sim \text{Binomial}(N_{ij}, p_{ij}) \quad (6)$$

$$\text{logit}(p_{ij}) \sim \text{Normal}(\alpha_j + \gamma_j A_i, \tau^{-1}) \quad (7)$$

$$A_i = X_i^T \beta \quad (8)$$

Here, α_j represents the degree of difficulty in league j , and γ_j represents the impact of ability on performance in league j . Accounting for league differences in this way enables us to compare, for example, a 0.400 OBP player in AAA ball to a 0.320 OBP player in the MLB. We constrain γ_j to be positive, so that a higher latent ability level corresponds to a higher probability of reaching base. We use $\tau^{-1} > 0$ to allow for additional player heterogeneity when modeling the p_{ij} . We include all of the sensorimotor variables in X_i with the exception of Go/No-Go, since it is highly correlated with



the Eye-Hand Coordination task and has limitations as a task [6]. We transform Depth Perception to the log scale as it is highly right-skewed and model the performance effects of age as linear. Diagnostics indicated that modeling age linearly fits the data reasonably well for all outcome models. We note that our findings were robust to both non-linear models for age and a maximum age threshold, mainly because age and sensorimotor tasks have weak associations in our sample (see also [9]). In addition to standardized age, X_i includes an indicator for catcher and an indicator for infielder. Hence, interpretations of all position coefficients are with respect to outfielders. Ultimately, we are interested in performing inference on the posterior distribution of β , which represents the impact of sensorimotor abilities on A_i . We use non-informative normal-gamma priors on β and τ ; see section **A.4.** for details.

Normal Response

SLG and FIP are long-run statistical averages over the number of at-bats and innings pitched, respectively. We present the model for SLG below; the model for FIP uses the same format. Let SLG_{ij} be the slugging percentage for player i in league j in N_{ij} at-bats. By the central limit theorem, as N_{ij} increases, the sampling distribution of SLG_{ij} approaches a normal distribution with some mean μ_{ij} and variance σ^2/N_{ij} . Because we only included player/league combinations such that $N_{ij} > 30$, the assumption of normality is plausible. We then specify a Bayesian multilevel model conditional on the slugging percentage ability parameters and league adjustment parameters. We have

$$SLG_{ij} \sim Normal\left(\mu_{ij}, \frac{\sigma^2}{N_{ij}}\right) \quad (10)$$

$$\mu_{ij} \sim Normal(\alpha_j + \gamma_j A_i, \tau^{-1}) \quad (11)$$

$$A_i = X_i^T \beta \quad (12)$$

The procedure for estimating this model is analogous to the binomial response case, but with an inverse-gamma prior distribution for σ^2 . The prior specifications are provided in section **A.4.**

2.3. Model Estimation

The models outlined in (7) - (12) are not identifiable since α_j , γ_j , and A_i are unknown and depend upon each other. We overcome this problem by imposing highly concentrated priors on α_j and γ_j , obtained by modeling the game statistics of all professional baseball players between 2012 and 2013 who played in multiple leagues. Details about this process are available in section **A.3.**

The posterior means of the league effect parameters α_j and γ_j obtained via the model of game statistics with all professional players are summarized in **Tables 3 and 4.** In particular, **Table 3** illustrates that there are two significant jumps in difficulty in professional baseball. There is a sizable increase in the quality of competition between Rookie baseball and non-rookie minor league baseball (A-AAA). In addition, there is an immense gap between AAA and the Major Leagues. Our model was unable to differentiate significantly between the non-rookie minor leagues. From **Table 4**, the impacts of ability are consistent across leagues, with the exception of the Major Leagues. With some statistics, such as OBP, latent ability matters less in the Major Leagues than it does in others. With others, such as FIP, it matters much more.



Attribute	Rookie	A	Adv. A	AA	AAA	MLB
logit ⁻¹ (OBP)	0.358	0.327	0.327	0.322	0.329	0.292
logit ⁻¹ (BB%)	0.108	0.093	0.096	0.093	0.089	0.071
logit ⁻¹ (K%)	0.170	0.188	0.184	0.192	0.195	0.232
SLG	0.432	0.384	0.379	0.376	0.397	0.351
FIP	3.013	3.517	3.377	3.613	3.782	4.279

Table 3: Posterior Means for α_j displaying the inverse-logit of the means for OBP, BB%, and K% for interpretability. In context, we project that an average professional player will obtain a 0.358 OBP in Rookie ball and a 0.292 OBP in the MLB

Attribute	Rookie	A	Adv. A	AA	AAA	MLB
OBP	0.110	0.118	0.109	0.101	0.103	0.060
BB%	0.316	0.304	0.300	0.320	0.305	0.275
K%	0.327	0.356	0.335	0.346	0.345	0.341
SLG	0.059	0.050	0.045	0.041	0.047	0.027
FIP	0.356	0.405	0.383	0.343	0.442	0.556

Table 4: Posterior Means for γ_j . Higher values indicate a higher relative impact of ability on the corresponding game statistic, given the model. These values should not be compared across statistics, since they are on different scales.

Once strong prior information on α_j and γ_j is obtained, we estimate the models detailed in equations (7) - (12), restricting our attention to the seasons of 141 batters and 111 pitchers in our sample with greater than 30 at-bats or innings pitched in each league. While it is reasonable to include data from all 149 batters when estimating the binomial response models, we elect to use the same player pool in all our models for consistency. To facilitate efficient Gibbs sampling and generate comparable coefficients, we standardize all variables in X_i with the exception of the position dummy variables. Although measurements of Depth Perception are missing for four batters and four pitchers, the missing values are sampled as part of the Gibbs sampler used to



estimate the model [16] with an independent standard normal prior placed on each of the missing values. We ran the model for three chains of 10,000 iterations after a 1000 iteration burn-in period, and validated it using Markov Chain Monte Carlo diagnostics and posterior predictive checks.

3. Results

To start off the analysis, we check to see if performance on the battery of sensorimotor tasks predicts on-field performance. In doing so, for each response variable, we fit two separate models: one with the sensorimotor tasks included and one with only age and position included as control variables. If sensorimotor abilities predict on-field performance, the full model should significantly outperform the reduced model. Following upon this, we report the individual coefficients for each of the models in which sensorimotor abilities added predictive power beyond the control variables.

3.1. WAIC

The Watanabe-Akaike Information Criterion (WAIC) is a useful way to compare two different Bayesian models of a particular response. It uses the log-posterior predictive density as the primary measure of accuracy, with a correction based upon the effective number of parameters in the model [17]. Asymptotically, it can be shown that WAIC approaches the results obtained via leave-one-out cross-validation [18]. For each of the five models, we use WAIC to compare the full model with the sensorimotor task results included in the design matrix to the reduced model that only accounts for position and age. If sensorimotor variables add predictive power above and beyond that of the control variables, then the WAIC of the full model should be lower than that of the reduced model.

Table 5 compares the WAIC of the full model to that of the reduced model for each of the five response variables. As indicated by the lower values for the Full, relative to the Reduced model, performance on the Nike Sensory Station tasks is predictive of OBP, BB%, and K%. However, sensorimotor abilities do not predict either SLG or FIP. We therefore present coefficient summaries for OBP, BB%, and K% in the sections below. Summaries for SLG and FIP can be found in section A.5.

	OBP	BB%	K%	SLG	FIP
Full Model	1210.8	1075.8	1276.4	403.4	363.8
Reduced Model	1226.4	1084.4	1284.6	403.1	361.9

Table 5: WAIC Model Comparison. Lower values for the full models relative to the reduced OBP, BB%, and K% models indicate that the added variables in the full models add meaningful predictive power.

3.2. Model Summaries

The posterior means, standard deviations, and 95% credible intervals for the coefficients β are presented in **Table 6** for the full OBP, BB%, K% models, and in **the Appendix** for SLG and FIP. The control covariates that are included in both the full and reduced models are indicated in the left sidebar. Variables for which 0 falls outside the 95% credible intervals are bolded. In general, bolded positive coefficients indicate that there is greater than 95% probability that the



sensorimotor ability measured in the task has an association with on-field performance. To illustrate the posterior tail probabilities, a heat map of the z-scored coefficients for OBP, BB% and K% is given in **Figure 2**.

		(A) On-Base Percentages				(B) Walk Rate				(C) Strikeout Rate			
		Mean	SD	2.5%	97.5%	Mean	SD	2.5%	97.5%	Mean	SD	2.5%	97.5%
Only Full Model	Visual Clarity	-0.24	0.17	-0.59	0.10	-0.15	0.10	-0.35	0.05	-0.08	0.07	-0.21	0.06
	Contrast Sensitivity	0.13	0.16	-0.18	0.45	0.04	0.09	-0.14	0.23	0.14	0.06	0.02	0.26
	Depth Perception	0.19	0.16	-0.12	0.50	0.21	0.10	0.02	0.40	-0.12	0.07	-0.26	0.02
	Near-Far Quickness	-0.02	0.15	-0.32	0.28	-0.05	0.09	-0.23	0.14	0.21	0.07	0.09	0.34
	Target Capture	0.15	0.16	-0.16	0.47	-0.16	0.09	-0.35	0.01	0.16	0.06	0.04	0.29
	Perception Span	0.64	0.17	0.31	0.99	0.15	0.10	-0.04	0.34	0.34	0.07	0.21	0.47
	Eye-Hand Coordination	0.22	0.17	-0.11	0.56	0.46	0.10	0.26	0.67	-0.19	0.07	-0.32	-0.06
	Reaction Time	0.21	0.17	-0.11	0.55	0.23	0.11	0.03	0.44	0.12	0.07	-0.02	0.26
Both	Age	0.66	0.17	0.34	1.00	0.53	0.09	0.36	0.71	0.22	0.06	0.09	0.34
	Infield	-0.53	0.31	-1.15	0.08	0.05	0.19	-0.33	0.43	0.65	0.13	0.40	0.91
	Catcher	-1.28	0.49	-2.25	-0.35	0.15	0.29	-0.40	0.72	0.26	0.19	-0.12	0.64
	Intercept	-0.13	0.23	-0.57	0.31	-0.84	0.14	-1.12	-0.57	-0.52	0.09	-0.71	-0.34

Table 6: Mean coefficients, standard deviations, and 95% credible intervals for each model variable are shown for (A) on-base percentage, (B) walk rate, and (C) strikeout rate. Values for which the 95% credible interval excludes zero are bolded.

From the OBP model results, we observe that performance on the Perception Span task, which measures the ability to remember and recreate visual patterns, is associated with an increased ability to reach base. Moreover, the size of the coefficient is comparable to that of age, a remarkable result considering it is well known that older players tend to perform better than younger players in professional baseball due to survivorship bias. For interpretation, suppose player X is a 23-year-old outfielder with completely average abilities as a professional baseball player. The model predicts his OBP in the MLB to be .292. We expect a similar player who scores one standard deviation higher on the Perception Span task to have an OBP of .300, a nontrivial difference. While the coefficients of the other tasks trend positive, there is simply not enough data to draw strong conclusions about them.

Furthermore, superior performance on the tasks that measure a player's ability to quickly identify and react to visual stimuli, Eye-Hand Coordination and Reaction Time, were found to be associated with an increased ability to draw walks. For example, our model predicts player X to obtain a walk rate of 7.1% in the MLB, but predicts a similar player with a one standard deviation superior score on the hand-eye coordination task to have a walk rate of 8.0%. On the other hand, superior performance on tasks that measure a player's spatial recognition and memory, such as Near-Far Quickness, Target Capture, and Perception Span, was found to be associated with an increased ability to avoid strikeouts. In context, our model predicts the strikeout rate of player X to be 23.2% in the MLB. A player similar to player X who scores one standard deviation better on the Perception



Span task is predicted to a strikeout rate of 21.2%. It is surprising that Eye-Hand Coordination was found to be significant in the opposite direction than we would expect a priori, which motivates further study.

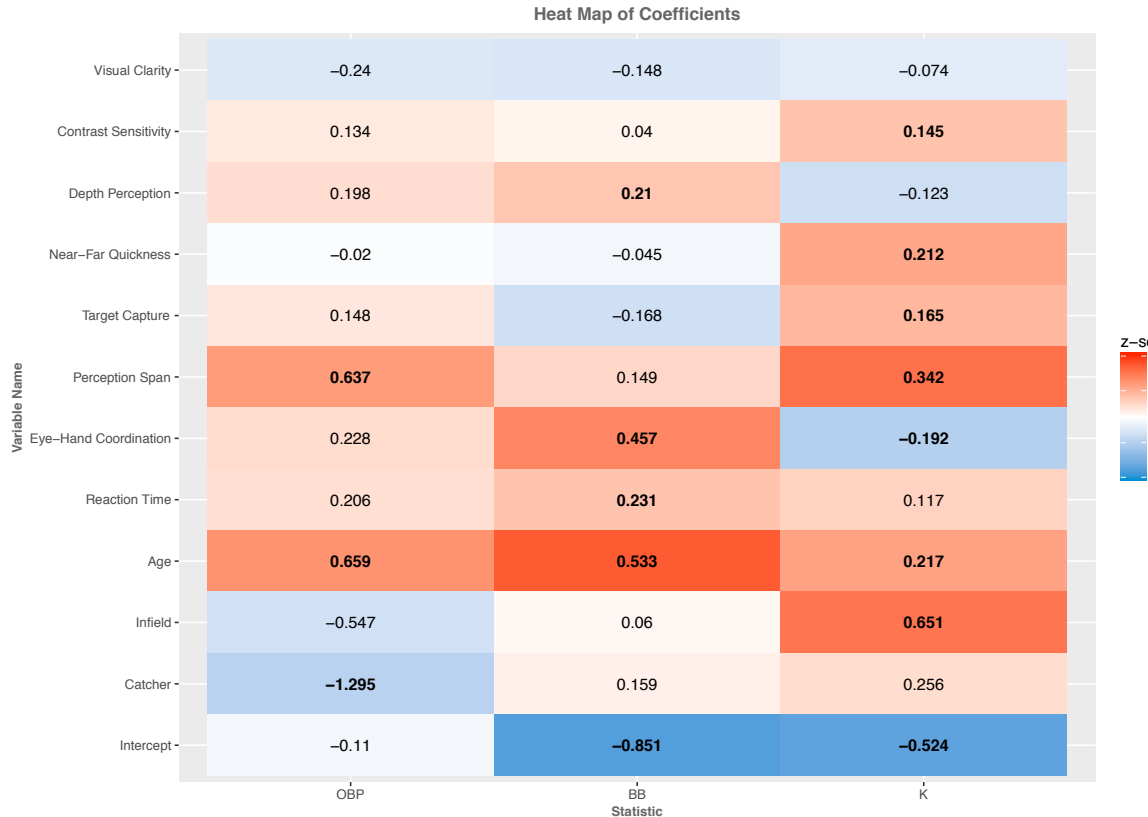


Figure 2: Heat map of β Coefficients. The darker the color, the closer the posterior tail probability gets to zero (indicating evidence of an association).

4. Discussion

The specific roles of vision, perception and motor control in interceptive sports such as baseball and cricket have been a hotly debated topic for years [3, 19, 20, 21]. In the current study, we shed new light on this debate by using real-world data collected from a large sports performance program launched by Nike Inc. Through Bayesian hierarchical latent variable modeling of the relationship between psychometric performance on the task battery and season-wide game statistics, we find that sensorimotor abilities predict on-base percentage, walk rate, and strikeout rate, but not slugging percentage or fielder-independent pitching.

The observation that better sensorimotor abilities generally correlate with better on-base percentage, walk rate, and strikeout rate is largely intuitive since it is expected that players draw on



these skills to project the location of the pitch through the strike zone and decide whether to swing or not. Conversely, the ability to hit for power, captured by slugging percentage, should have more to do with strength, bat speed, and swing plane than sensorimotor abilities. Pitchers rely on a strong arm, consistent mechanics, and a varied repertoire to prevent runs, attributes that are superficially unrelated to sensorimotor abilities.

Among the individual tasks tested, Perception Span shows the strongest relationship with on-field performance, with better scores strongly associated with both increased on-base percentages and reduced strikeout rates. In addition, performance on the Perception Span task exhibits some association with both higher walk rates and increased slugging percentages, though the evidence is not conclusive. This task measures the ability to remember and recreate visual patterns and may reflect visual recognition abilities that have previously been tied to batting performance in small samples of players (N=20) tested with Tachistoscopic methods [22] and in a conference paper reporting relationships in collegiate players [23].

A number of other tasks correlate with higher walk rates, including Depth Perception, Eye-Hand Coordination and Reaction Times. The observation that Eye-Hand Coordination and Reaction Time are positively correlated with walk rate indicates that the ability to quickly react to visual stimuli is highly influential in a player's ability to draw walks. The positive relationship with Depth Perception supports previous findings indicating that binocular vision contributes to precisely projecting the location of a pitched baseball [24]. Furthermore, past research comparing pitchers and hitters on the Sensory Station battery found better performance for professional hitters, relative to pitchers, on both the Visual Clarity and Depth Perception tasks [9], suggesting that better depth disparity differentiates highly experienced athletes who bat from those who pitch.

The model linking sensorimotor abilities to strikeout rates offers a mixed view of the relative contributions of sensorimotor skills towards avoiding strikes. The pattern of results indicates that better performance on the Perception Span, Near-Far Quickness, Target Capture, and Contrast Sensitivity tasks is associated with an increased ability to avoid strikeouts. However, it is surprising that worse Eye-Hand Coordination is associated with reduced strikeout rates, though it has a relatively weak coefficient.

In light of the current findings, it is worth noting several strengths and weaknesses in the approach. First, this dataset reflects one of the largest samples of high-level baseball players tested on a consistent battery of psychometric tasks. These tasks were presented with video instructions and conducted by trained and certified administrators, providing some assurance towards data quality. Further, the latent approach to modeling league heterogeneity offers a systematic means by which to incorporate data from multiple leagues, while also scaling production in each league to accurately reflect the relative difficulty of that league in that year. Nonetheless, it is important to note that while the individual tasks in the battery have been identified as important abilities for sports performance [4] the choice to include multiple measures in the battery comes with a tradeoff of fewer trials (and less sensitivity) for each measure. On the baseball side, the data shared with the research team limited the scope of potential research directions, preventing us from incorporating park-specific factors and more granular, robust metrics of player performance.

4.1. Future Work

Given the incredible amounts of data that technological tools such TrackMan and StatCast generate in professional baseball, this study only scratches the surface of the informational gold-mine that can be harvested by teams, researchers, and vision centers who want to understand the



relationship between sensorimotor skills and on-field performance. Below, we present a few ideas for follow-up studies that can be conducted by teams or researchers under a secondary-data protocol.

One interpretation of the strong relationship between Perception Span and batting performance is that the ability to store pitches in spatial working memory, and subsequently recognize them, helps batters avoid strikeouts and reach base more frequently. There may be evidence for this empirically, since pitchers obtain the highest strikeout rates and allow the lowest on-base percentage when seeing batters for the first time. Each time a batter faces a pitcher, his on-base percentage improves and strikeout rate declines [25], in part because he has “seen” the pitcher’s repertoire before and filed it away into memory, making for easier recollection and recognition during subsequent meetings. Future research may examine whether players with high scores on the Perception Span task perform better against pitchers during the second and third times through the order, above and beyond the improvement expected of them. To do this, at-bat level data will be needed, rather than aggregated season statistics.

Walk rate is an imperfect metric for plate discipline because it captures a batter’s underlying approach to the plate, in addition to measuring a batter’s ability to differentiate between balls and strikes. There’s no question that a batter’s vision plays an integral role in his decision to swing. In 2008, Carlos Santana, one of the most disciplined first basemen in the MLB, independently began to practice pitch visualization, imagining a ring between the pitcher’s mound and home-plate. If the ball traveled through the ring, he would make the decision to swing and otherwise hold back [26]. Perhaps analyzing pitch-level data can clarify the specific sensorimotor skills that help batters like Santana make optimal decisions.

One important attribute not considered in this study is a player’s defensive ability. With Statcast, which tracks the location of every player on the field at 25 frames per second, teams have the ability to track the route efficiency and first step of outfielders as they attempt to catch fly balls. An analysis of the relationship between sensorimotor abilities and route efficiency could provide unique insight into player defense, particularly for outfielders.

If the present and future results speak to underlying building blocks of baseball expertise, how can they be used to improve baseball performance? This question lies at the heart of efforts to implement “sports vision training” programs [27, 28] based on the notion that practice with demanding visual, perceptual, cognitive, or oculomotor tasks can improve the ability to process and respond to what is seen, thereby improving athlete performance. The literature has examined training techniques that target anticipation and decision-making abilities of athletes [29], as well as new digital technologies that train general visual, perceptual and cognitive skills critical for sporting performance [12, 30, 31]. Ultimately, the ability to determine which visual and motor characteristics are related to performance will focus research on specific training programs, enabling athletes to make the most of their system physiologies. The current findings are an important step in this direction.



References

- [1] Bahill, A.T. and T. LaRitz, *Why can't batters keep their eyes on the ball?* American Scientist, 1984. **72**(3): p. 249-253.
- [2] Molia, L.M., S.E. Rubin, and N. Kohn, *Assessment of stereopsis in college baseball pitchers and batters.* J aapos, 1998. **2**(2): p. 86-90.
- [3] Epstein, D., *The sports gene: Inside the science of extraordinary athletic performance.* 2013: Penguin.
- [4] Erickson, G.B., et al., *Reliability of a computer-based system for measuring visual performance skills.* Optometry, 2011. **82**(9): p. 528-42.
- [5] Gilrein, T., *Reliable change indices of visual and sensory performance measures, in Department of Exercise and Sports Science.* 2014, University of North Carolina, Chapel Hill.
- [6] Wang, L., et al., *Mapping the structure of perceptual and visual-motor abilities in healthy young adults.* Acta Psychol (Amst), 2015. **157**: p. 74-84.
- [7] Krasich, K., et al., *Sensorimotor learning in a computerized athletic training battery.* Journal of Motor Behavior, 2016. **48**(5): p. 401-412.
- [8] Appelbaum, L.G., et al., *The Effects of Sports Vision Training on Sensorimotor Abilities in Collegiate Softball Athletes.* Athletic Training & Sports Health Care, 2016.
- [9] Klemish, D., et al., *Visual Abilities Distinguish Pitchers from Hitters in Professional Baseball.* . Journal of Sports Sciences, 2017.
- [10] Harpham, J.A., et al., *The effect of visual and sensory performance on head impact biomechanics in college football players.* Annals of Biomedical Engineering, 2014. **42**(1): p. 1-10.
- [11] Poltavski, D. and D. Biberdorff, *The role of visual perception measures used in sports vision programs in predicting actual game performance in Division I collegiate hockey players.* Journal of Sports Science, 2014.
- [12] Appelbaum, L.G. and G.B. Erickson, *Sports vision training: a review of the state-of-the-art in digital training techniques.* International Review of Sport and Exercise Psychology, 2016. Published on line, Nov 23, 2016: p. 1-30.
- [13] DHHS, *Secretary's Advisory Committee on Human Research Protections. Attachment A: human subjects research implications of "big data" studies,* U.S.D.o.H.a.H. Services, Editor. 2015.
- [14] FanGraphs. *FanGraphs.com.* 2017 [cited 2017 March 2]; Available from: <http://www.fangraphs.com/>.
- [15] Gelman, A. and J. Hill, *Data analysis using regression and multilevel/hierarchical models.* 2006: Cambridge University Press.
- [16] Plummer, M., *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling.* 2003.
- [17] Gelman, A., J. Hwang, and A. Vehtari, *Understanding predictive information criteria for Bayesian models.* . Statistics and Computing, 2013. **24**(6): p. 997-1016.
- [18] Vehtari, A., A. Gelman, and J. Gabry, *Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC.* . Statistics and Computing, 2016. **27**(5): p. 1413-1432.
- [19] Elmurr, P., *The relationship of vision and skilled movement-a general review using cricket batting.* Eye Contact Lens, 2011. **37**(3): p. 164-6.



- [20] Miller, B.T. and W.C. Clapp, *From vision to decision: the role of visual attention in elite sports performance*. Eye Contact Lens, 2011. **37**(3): p. 131-9.
- [21] Laby, D.M., *Sports Vision*, in *Focal Points*. 2014, American Academy of Ophthalmology: San Francisco.
- [22] Reichow, A.W., K.E. Garchow, and R.Y. Baird, *Do scores on a tachistoscope test correlate with baseball batting averages?* Eye Contact Lens, 2011. **37**(3): p. 123-6.
- [23] Szymanski, D., Light, T], Voss, Z], and Greenwood, M., *Relationships between vision performance scores and offensive statistics of college baseball players.* , in *American College of Sports Medicine, Southsat Regional Chapter meeting*. 2015, Medicine and Science in Sports and Exercise: Jacksonville, Florida.
- [24] Hofeldt, A.J., F.B. Hoefle, and B. Bonafede, *Baseball hitting, binocular vision, and the Pulfrich phenomenon*. Arch Ophthalmol, 1996. **114**(12): p. 1490-4.
- [25] Lichtman, M. *Baseball ProGUESTus: Pitch Types and the Times Through the Order Penalty*. 2013 [cited 2017 May 11]; Available from: <http://www.baseballprospectus.com/article.php?articleid=22235>.
- [26] Sawchik, T. (2017). Carlos Santana's honed batting eye is valuable but how it might age (and affect the Indians' plans) is the question. The Athletic.
- [27] Deveau, J., D.J. Ozer, and A.R. Seitz, *Improved vision and on-field performance in baseball through perceptual learning*. Curr Biol, 2014. **24**(4): p. R146-7.
- [28] Erickson, G.B., *Sports vision: vision care for the enhancement of sports performance*. 2007, St. Louis, MO: Butterworth-Heinemann.
- [29] Broadbent, D.P., et al., *Perceptual-cognitive skill training and its transfer to expert performance in the field: future research directions*. Eur J Sport Sci, 2015. **15**(4): p. 322-31.
- [30] Appelbaum, L.G., et al., *Stroboscopic visual training improves information encoding in short-term memory*. Atten Percept Psychophys, 2012. **74**(8): p. 1681-91.
- [31] Appelbaum, L.G., et al., *Improved Visual Cognition through Stroboscopic Training*. Front Psychol, 2011. **2**: p. 276.



Appendix

A.1. Sensorimotor Assessments

The Sensory Stations consist of a battery of nine computerized sensorimotor tasks, each designed to evaluate a specific facet of a participant's visual-motor abilities. The first five tasks were completed using a handheld Apple iPod Touch, standing 4.9 m from the station. The last four tasks were completed at arm's length from the touchscreen monitor. Four of the tasks – Visual Clarity, Contrast Sensitivity, Depth Perception, and Target Capture – operated on staircase schedules in which subsequent stimulus difficulty increased following a correct response and decreased following an incorrect response. For these tasks, scores were calculated as the final step according to response accuracy on the staircase schedule. All tasks were preceded by video instructions. Procedures and descriptions for each task are provided below, and detailed descriptions can be found in [A1] and [A3].

The **Visual Clarity** task measures visual acuity for fine details at a distance using a black Landolt ring – an incomplete ring with a small gap oriented in one of the four cardinal directions. Participants were asked to swipe on the iPod in the direction that corresponded to the orientation of the gap in the ring. The task was completed in three separate rounds: one with an occluder covering the right eye, then the left, then a final round with both eyes uncovered. Visual Clarity scores were taken as the average of these three conditions.

The **Contrast Sensitivity** task measures the minimum resolvable difference in contrast at a distance. Participants were presented with four black rings on a light gray background and asked to indicate which ring contained a pattern of dark gray concentric circles by swiping on the iPod in the direction corresponding to the patterned ring.

The **Depth Perception** task measures how quickly and accurately participants are able to detect differences in depth at a distance using liquid crystal glasses. Here, four black rings were presented and participants were asked to swipe in the direction of the ring that appeared to have depth. The task was completed three times: once facing towards the screen, once facing to the left and looking over the right shoulder, and once facing right and looking over the left shoulder. Depth Perception scores were taken as the average of these three conditions.

The **Near-Far Quickness** task measures the number of near and far targets that can be correctly reported in 30 seconds. Participants aligned the top of the iPod with the bottom edge of the large monitor then swiped in the direction of the gap in the Landolt ring that appeared on either the iPod or larger monitor screen. Participants were instructed to respond as quickly as possible, and the ring only moved from one screen to another following a correct response. Participants continued to respond until they answered correctly or ran out of time. Near-Far Quickness scores were the total number of correct responses made in 30-seconds.

The **Target Capture** task measures the speed at which participants can shift attention and recognize peripheral targets. A small black Landolt ring was briefly presented in one of the four corners of the monitor, and participants were asked to swipe on the iPod in the direction corresponding to the gap in the ring. Following a correct answer, the ring was presented for a shorter duration, and for a longer duration following an incorrect answer, per the staircase



procedure. Because this task was performed on a duration staircase, the final accuracy step reflected the minimum stimulus duration according to accuracy on the staircase schedule.

The **Perception Span** task measures the capacity of spatial working memory. As participants stood at arm's length from the monitor, a grid of empty black circles was presented, and a subset was filled briefly with green dots that disappeared after 100 milliseconds. Participants were asked to recreate the pattern on each trial by touching the circles that had previously contained the green dots. There were eleven total possible pseudo-randomized trials with increased grid sizes and increasing number of green dots presented at each level. Perception Span scores were computed as the total number of correctly identified dots minus the number of missed or falsely identified dots across all of the trials.

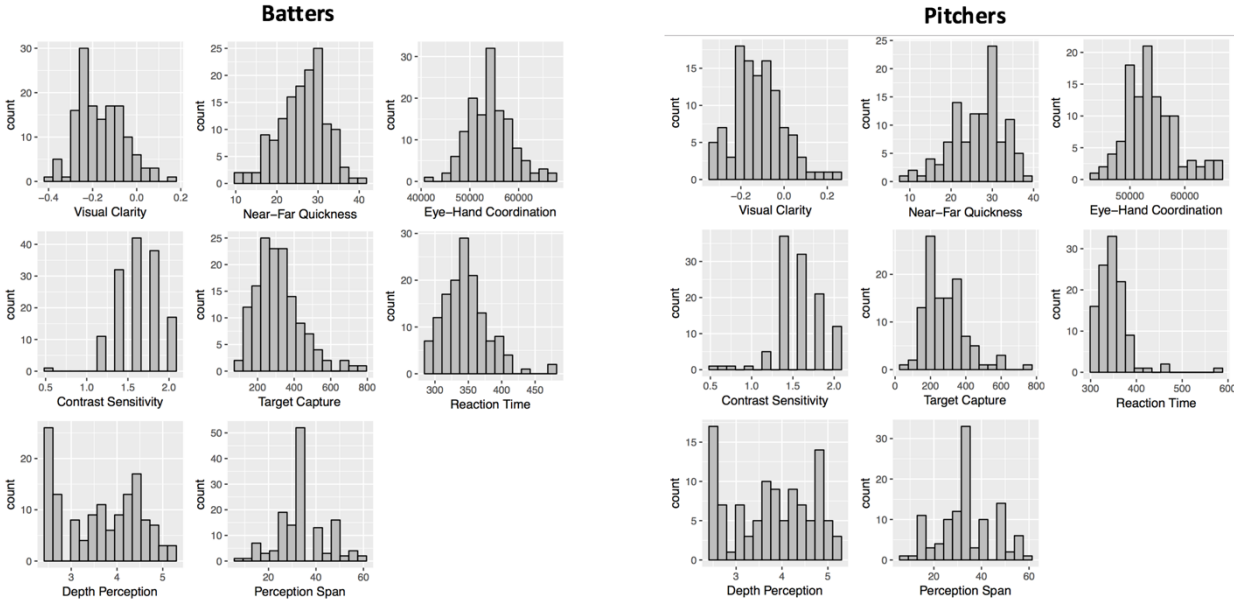
The **Eye Hand Coordination** task measures the speed at which participants can make visually-guided hand responses to rapidly changing targets. A grid of 48 evenly spaced black rings was presented on the screen. When a green dot appeared in one of the rings, participants touched the dot as quickly as possible. The dot then relocated to another ring for a total succession of 96 dots. The score for Eye Hand Coordination was the total time it took to complete the sequence.

The **Go/No-Go** task measures the ability to execute and inhibit visually guided hand responses in the presence of "go" and "no-go" stimuli. Similar to the previous task, a grid of 48 rings was presented; however, in this task the dots could appear either green or red. Participants tried to touch the green dots as quickly as possible while avoiding red dots. 96 dots were presented for 500 milliseconds each before disappearing, and the total score was calculated as the number of green dots touched minus the number of red dots touched.

The **Response Time** task measures how quickly participants react and respond to a simple visual stimulus. Two rings were shown on each side of the large monitor. Participants began with their dominant hand in the "starting" ring, while their body was oriented in front of the "landing" ring on the opposite side of the screen. When the landing ring turned green, participants moved their hand from the starting ring to the landing ring as quickly and accurately as possible. A total of seven separate trials were completed, and participants had the opportunity to repeat up to two of these trials if any were slower than two standard deviations from the mean. Response Time scores were taken as the average of the seven best trials.



A.2. Performance Distributions for the Sensory Station tasks



Histograms of behavioral performance on the Sensory Station tasks for (L) batters and (R) pitchers.

A.3. League Effect Models

To estimate α_j and γ_j , we first scrape publicly available minor and major league play-by-play data between 2012 and 2013 from MLB.com using the pitchRx package in the R language [A2]. We collate the statistics of all players (1695 batters, 1053 pitchers) who played in multiple leagues during that period. By examining the difference in player performance between the leagues, we quantify the degree of difficulty of each league with a separate Bayesian model. For example, if the Major League is more difficult than AAA, we should expect a player who plays in both leagues to register a lower on-base percentage in the Major League than in AAA.

For each of the five game statistic variables, we estimate the corresponding model detailed in Eqs.(7) - (12), using the data of all players who played in multiple leagues between 2012 and 2013. Since we do not have sensorimotor measurements for these players, we instead place a standard normal prior on each A_i . Because we assume that increased ability corresponds to improved performance for each game statistic, we impose a half-Cauchy prior on γ_j , as well as conjugate Normal/Gamma priors for α_j and τ .

For the initial on-base percentage model, we use the following specification and priors.

$$\begin{aligned}
 OB_{ij} &\sim \text{Binomial}(N_{ij}, p_{ij}) & \text{logit}(p_{ij}) &\sim \text{Normal}(\alpha_j + \gamma_j A_i, \tau^{-1}) & A_i &\sim \text{Normal}(0, 1) \\
 \alpha_j &\sim \text{Normal}(0, \tau_\alpha^{-1}) & \tau_\alpha &\sim \text{TN}(0, 1, 0, \infty) & & \\
 \gamma_j &\sim \text{TN}(0, \tau_\gamma^{-1}, 0, \infty) & \tau_\gamma &\sim \text{TN}(0, 1, 0, \infty) & \tau &\sim \text{Gamma}(0.5, 0.5)
 \end{aligned}$$



where $TN(\mu, \sigma, a, b)$ is the normal distribution truncated to a lower bound a and upper bound b . The initial models for BB% and K% models have similar specifications, with the exception that γ_j in the K% model is bounded above by zero, since players with higher ability strike out less frequently.

For the initial slugging percentage model, we use the following specification and priors.

$$\begin{aligned}
 SLG_{ij} &\sim Normal\left(\mu_{ij}, \frac{\sigma^2}{N_{ij}}\right) & \mu_{ij} &\sim Normal(\alpha_j + \gamma_j A_i, \tau^{-1}) & A_i &\sim Normal(0, 1) \\
 \alpha_j &\sim Normal(0, \tau_\alpha^{-1}) & \tau_\alpha &\sim TN(0, 1, 0, \infty) & 1/\sigma^2 &\sim Gamma(0.5, 0.5) \\
 \gamma_j &\sim TN(0, \tau_\gamma^{-1}, 0, \infty) & \tau_\gamma &\sim TN(0, 1, 0, \infty) & \tau &\sim Gamma(0.5, 0.5).
 \end{aligned}$$

The specification for the FIP model is the same as that for the SLG model, with the exception that γ_j in the FIP model is bounded above by zero, since players with higher ability record lower values for FIP.

For each model, we draw ten thousand samples of the parameters from the joint posterior distribution via Gibbs sampling. We compute the posterior means and variances for all parameters, including α_j and γ_j . These means and variances are used in the prior distributions for the final models, as described in **A.4**.

A.4. Specifications for Final Models

As mentioned in the main body of the article and in **A.3.**, we first fit an initial model to obtain concentrated priors for α_j , γ_j , τ , and σ^2 . We use the posterior means and variances to form prior distributions in the final models used to assess the relationship between sensorimotor abilities and on-field performance. For any initial model, let the $\hat{\tau}$ and τ_{SD} be the posterior mean and standard deviation of τ , respectively, computed as described in **A.3**. We use analogous notation for the posterior means and standard deviations for the other parameters. We emphasize that each outcome variable has its own set of parameters, although we use a common notation for convenience.

The final on-base percentage model used to produce the results in **Table 6**, including all prior distributions, is as follows:

$$\begin{aligned}
 OB_{ij} &\sim Binomial(N_{ij}, p_{ij}) & \text{logit}(p_{ij}) &\sim Normal(\alpha_j + \gamma_j A_i, \tau^{-1}) & A_i &= X\beta \\
 \alpha_j &\sim Normal(\hat{\alpha}_j, \alpha_{SDj}^2) & \gamma_j &\sim TN(\hat{\gamma}_j, \gamma_{SDj}^2, 0, \infty) & \beta_j &\sim Normal(0, 10^2), \forall j \\
 & & \tau &\sim TN(\hat{\tau}, \tau_{SD}^2, 0, \infty) & X_{i,miss} &\sim Normal(0, 1)
 \end{aligned}$$

The specifications for the BB% and K% models are identical, with the exception that γ_j in the K% model is bounded above by zero, since players with higher ability strike out less frequently.

The final slugging percentage model used to produce the results in **Table 6**, including all prior distributions, is as follows.

$$\begin{aligned}
 SLG_{ij} &\sim Normal\left(\mu_{ij}, \frac{\sigma^2}{N_{ij}}\right) & \mu_{ij} &\sim Normal(\alpha_j + \gamma_j A_i, \tau^{-1}) & A_i &= X\beta
 \end{aligned}$$



$$\alpha_j \sim \text{Normal}(\hat{\alpha}_j, \alpha_{SDj}^2) \quad \gamma_j \sim \text{TN}(\hat{\gamma}_j, \gamma_{SDj}^2, 0, \infty) \quad \beta_j \sim \text{Normal}(0, 10^2), \forall j$$

$$\sigma \sim \text{TN}(\hat{\sigma}, \sigma_{SD}^2, 0, \infty) \quad \tau \sim \text{TN}(\hat{\tau}, \tau_{SD}^2, 0, \infty) \quad X_{i,miss} \sim \text{Normal}(0,1)$$

The specification for the FIP model is the same as that for the SLG model, with the exception that γ_j in the FIP model is bounded above by zero, since players with higher ability record lower values for FIP.

A.5. SLG and FIP Model Summaries

		(A) Slugging Percentage				(B) Fielder Independent Pitching			
		Mean	SD	2.5%	97.5%	Mean	SD	2.5%	97.5%
Full Model	Visual Clarity	-0.01	0.15	-0.30	0.28	-0.24	0.25	-0.71	0.26
	Contrast Sensitivity	-0.14	0.14	-0.41	0.15	0.07	0.24	-0.40	0.54
	Depth Perception	0.18	0.14	-0.08	0.45	-0.68	0.22	-1.12	-0.26
	Near-Far Quickness	0.00	0.14	-0.26	0.27	0.25	0.23	-0.20	0.70
	Target Capture	0.15	0.14	-0.12	0.43	-0.11	0.21	-0.52	0.29
	Perception Span	0.29	0.14	0.01	0.58	0.21	0.23	-0.25	0.66
	Eye-Hand Coordination	0.05	0.15	-0.25	0.34	-0.18	0.20	-0.58	0.20
	Reaction Time	0.09	0.15	-0.21	0.40	-0.07	0.20	-0.47	0.34
Reduced	Age	0.44	0.15	0.15	0.72	0.55	0.21	0.15	0.96
	Infield	-0.60	0.28	-1.13	-0.05				
	Catcher	-1.40	0.41	-2.21	-0.60				
	Intercept	0.07	0.20	-0.33	0.45	-0.32	0.21	-0.74	0.07

Mean coefficients, standard deviations, and 95% credible intervals for each model variable are shown for (A) slugging percentage and (B) fielder-independent pitching. Values for which the 95% credible interval excludes zero are bolded.

A.6. Supplemental References

[A1] Erickson, G. B., Citek, K., Cove, M., Wilczek, J., Linster, C., Bjarnason, B., & Langemo, N. (2011). Reliability of a computer-based system for measuring visual performance skills. *Optometry*, 82(9), 528-542. doi:10.1016/j.optm.2011.01.012

[A2] Sievert, C. (2015). pitchRx: Tools for Harnessing 'MLBAM' 'Gameday' Data and Visualizing 'pitchfx' (Version R package version 1.8.2.).

[A3] Wang, L., Krasich, K., Bel-Bahar, T., Hughes, L., Mitroff, S. R., & Appelbaum, L. G. (2015). Mapping the structure of perceptual and visual-motor abilities in healthy young adults. *Acta Psychol (Amst)*, 157, 74-84. doi:10.1016/j.actpsy.2015.02.005